HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Molecular and Cellular Biology

have examined a dissertation entitled

"Genome-scale DNA Methylation analysis in distinct pluripotent stem cells"

presented by Jamie Orme Webster,

candidate for the degree of Doctor of Philosophy and hereby

certify that it is worthy of acceptance.

Signature ______________________________

Typed name: Prof. Catherine Dulac

Signature ______________________________

Typed name: Prof. John Rinn

Signature ______________________________

Typed name: Prof. Qiao Zhou

Date: April 16, 2014

Genome-scale DNA Methylation analysis in distinct pluripotent stem cells

A dissertation presented

by

Jamie Orme Webster

to

The Molecular and Cellular Biology (MCB) Department

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biology

Harvard University

Cambridge, Massachusetts

April 2014

Dissertation Advisor: Professor Alexander Meissner Jamie Webster

Genome-scale DNA Methylation analysis in distinct pluripotent stem cells.

Abstract

Mouse embryonic stem cells (ESCs) and embryonic germ cells (EGCs) are both pluripotent cell types derived from distinct developmental stages. Despite sharing many characteristics, some differences have been reported between the two cell types. For instance, EGCs generally lack of DNA methylation at imprinted regions and they have been suggested to display overall global hypomethylation. These characteristics have always been regarded as reflective of their derivation origin from primordial germ cells (PGCs), whereas ESCs are derived from preimplantation blastocysts. A large set of newly derived, genetically matched ESC and EGC lines has enabled us to clarify several of these observations at genome-wide scale. DNA methylation analysis of ESCs and EGCs, in particular lines derived from both sexes demonstrates that the DNA methylation profiles of female ESCs and EGCs are comparable low, while the male cell lines all share a more hypermethylated genome. The observed hypomethylation is not restricted to promoters but is also evident in many retroelements including the regions often regarded as resistant to demethylation. Mechanistically it appears that the two active X chromosomes in female cells are involved. Specifically, we find that the X-linked gene Dusp9 may play a central role in the regulation of DNA methylation via the inhibition of Erk signaling. Taken together, our results provide novel mechanistic insights as well as clarify that it is not the EGC identity that caused the global hypomethylation in prior studies, but rather the fact that the lines in those experiments happened to be female EGCs.

Table Of Contents

1.0 INTRODUCTION

2.0 RESULTS & DISCUSSION

3.0 ADDITIONAL MATERIAL

ACKNOWLEDGEMENTS

I would like to thank my committee members for taking the time to read this thesis. Thanks to fellow lab members and thanks to Zachary Smith and Julie Donaghey for proof reading. Special Thanks to Michael Ziller and Kendell Clement for all their help with the processing of sequencing data. Thanks to Michael Lawrence for all of his help. I would also like to thank my adviser Alex Meissner for his support throughout this process. Most of all I would like to thank my wife and daughter for everything else. This is dedicated to the memory of my father, William Webster.

List of Figures

INTRODUCTION

Origins of Pluripotency

The study of pluripotent stem cells dates back to the 1950s, with the discovery of embryonal carcinoma cells (ECCs) derived from teratocarcinomas in mice. These teratomas were found to consist of two populations of cells, an undifferentiated ECC component and the differentiated component featuring all three germ layers (Stevens and Little, 1954). The isolated ECCs were found to be capable of causing teratomas when serially transplanted in mice, due to the ECCs "stem cell" function, allowing it to generate all components of the tumor (Stevens and Little, 1954). It wasn't until a decade later that researchers were able to show that these ECCs were capable of a broad range of plasticity with regards to their differentiation capabilities in addition to maintaining unlimited self –renewal (Kleinsmith and Pierce, 1964). The ability to stably culture ECCs *in vitro* allowed them to be used as a powerful model of mammalian development (Kahan and Ephrussi, 1970). This capability of ECCs to mimic some normal developmental processes would eventually lead to the modern day field of stem cell biology.

From ICM to ESC

In mammalian development life begins with a single cell, the zygote. This cell proceeds through cleavage divisions to create 2-cell, 4-cell embryos, which initially maintain totipotency, which is the ability to generate all cell types. Around the morula stage (16-32 cells) the very first cell fate choice occurs and fate becomes more restricted. The outer layer of cells adopts a trophectoderm fate, while the inner cell mass (ICM) of the so-called blastocyst remains undifferentiated, but only pluripotent. The trophectoderm will eventually form the extra-embryonic tissue, while the ICM will give rise to all cells of the organism (Johnson and Ziomek, 1981).

In mice the blastocyst stage occurs at 3.5 days post coitum (dpc). Notably, embryos at the blastocysts can be readily removed from the uterus and explanted in appropriate culture conditions onto mitotically inactivated mouse embryonic fibroblasts (MEFs), or feeder cells as they are more commonly known, to give rise to stable cell lines, called embryonic stem cells (ESCs). ESCs were first isolated and cultured in 1981 (Martin, 1981). Interestingly the early culture required media conditioned by ECCs, indicating relevant similarities between the two cell types. While similar to the ECCs in many ways, including their self-renewing and differentiation capabilities, the ESCs provided an added benefit of a stable karyotype, unlike the ECCs which exhibited genomic abnormalities (Bradley et al., 1984).

To identify factors necessary for maintaining ESCs in culture, ECC conditioned media was fractionated. This systematic approach identified leukemia inhibitory factor (LIF) as one of the key players in maintaining a pluripotent state in ESCs (Smith et al., 1988). LIF functions in maintaining pluripotency via the JAK/STAT pathway (Fukada et al., 1996), promoting self-renewal while inhibiting differentiation. LIF, however, is only able to maintain pluripotency in the presence of serum (usually fetal bovine serum [FBS]). This was later found to be due to the presence of a TGFβ_ superfamily of proteins known as bone morphogenetic proteins (BMPs) in the serum. BMPs facilitate pluripotency by inducing inhibitor of differentiation (id) genes (Ying et al., 2003). ESCs can still be cultured with LIF in the absence of serum when supplemented with BMP4. More recently, it was shown that two inhibitors (2i) of the signaling kinases Mek and GSK3beta are also sufficient to maintain ESC undifferentiated in the absence of serum and Lif. This suggests inhibition of phosphorylated ERK is required to maintain pluripotency and that LIF/Stat signaling may be working downstream of this (Ying et al., 2008).

Despite being an *in vitro* artifact, ESCs share several features with their *in vivo* counterparts, the ICM cells. The four main pluripotency factors, Oct4, Sox2, Klf4, and c-myc, are found in both the ICM and cultured ESCs (Surani et al., 2007; Toyooka et al., 2008; Li et al., 2012). Several core transcription factors (TFs) including Oct4, Sox2 and Nanog cooperate to maintain the undifferentiated state of the ESCs and form a complex regulatory circuit (Rodda, 2005). Oct4 is necessary to activate protein-coding genes and non-coding RNAs required for pluripotency (Herr and Cleary, 1995). Additionally, Oct4 associates with components of the polycomb group complex (PRC) to repress differentiation programs. In the absence of Oct4 embryos fail to thrive and die around the time of implantation (Nichols et al., 1998). Another trait that female murine ESCs share with their ICM origin is the state of their X chromosomes. Female cells contain two X chromosomes, compared to the male cells' one (Lyon, 1962) and generally compensate for this, during normal development by inactivating one of them (Xi). Insights into this very complex process continue to emerge, but it is well understood that transcription of a non-coding RNA (ncRNA) called Xist is central to the mechanism. Transcription of this ncRNA initiates coating of the Xi chromosome in *cis*. This in turn is believed to recruit other factors, which eventually lead to heterochromatization of the Xi (Senner and Brockdorff, 2009). In contrast to somatic cells, in the ICM and female mouse ESCs both X chromosomes remain active (Barakat and Gribnau, 2010; Monk, 1981).

From Epiblast to EpiSC

While ESCs are derived from the pre-implantation blastocyst at 3.5 dpc, it is possible to isolate another group of pluripotent stem cells from the post-implantation epiblast at 5.5-6.5 dpc known as Epiblast Stem Cells (EpiSCs) (Najm et al., 20011). As previously mentioned, the pre-implantation blastocyst consists of the ICM and the outer layer of trophectoderm cells. After implantation a second cell fate decision occurs; the ICM gives rise to the epiblast and primitive endoderm (Takoaka and Hamada, 2012). Epiblast cells can be identified by their expression of pluripotency associated factors including Oct4 and Nanog, which are absent from the primitive endoderm, that shows expression of Gata4 and Gata6 instead. The ICM shows already a distinct pattern of Nanog expression indicating heterogeneity in this cell population that may be linked to this downstream fate (Chazaud et al., 2006). Once the epiblast is formed at 5.5 dpc, it can be isolated and in appropriate culture conditions give rise to EpiSCs (Brons et al., 2007, Najm et al. 20011)). Similar to pluripotent ESCs, EpiSCs can be differentiated *in vitro*, giving rise to all three germ layers. Additionally, EpiSCs may be injected into immuno-compromised mice and form teratomas showing multi-lineage potential (Brons et al., 2007). However there are some major differences between these two cell types. The first and most apparent difference is the morphology of the cell colonies in culture. ESCs colonies have a dome-shape while EpiSCs tend to have a flatter, more pancake like appearance. Unlike female ESCs, which maintain both Xs in an active state, EpiSCs have already randomly selected an Xi consistent with their later developmental point of derivation (Hayashi and Surani, 2009). Moreover they display key differences in the requirements for signaling pathways required for stable maintenance. ESCs are derived in media conditioned with LIF and BMP, whereas EpiSCs cannot be propagated under these conditions. Instead EpiSC derivation is dependent on Activin/Nodal signaling. One regulatory component is that Activin/Nodal signaling upregulates Nanog, which in turn prevents the differentiation into neuroectoderm by inhibiting FGF signaling. Additionally, Nanog inhibits the transcriptional activity of the Smad2/3 proteins, responsible for mesendoderm differentiation. Interestingly the up-regulation of Smad2/3 is caused by Activin/Nodal signaling. Nanog, therefore, appears to help maintain pluripotency in the EpiSC by maintaining the balance between ectoderm and mesendoderm signaling (Vallier et al., 2009). This ability of Nanog to prevent differentiation is also seen in ESCs and the ICM. However, in those instances, it acts to prevent primitive endoderm differentiation (Hamazaki et al., 2004; Mitsui et al., 2003).

While pluripotency refers to a cell's potential to differentiate into all cell types, it should be made clear that there are differences among pluripotent populations including ESCs and EpiSCs, *in vitro* EpiSCs, much like ESCs, can be easily differentiated into multiple lineages and under go many passages without losing their self-renewal capabilities, *in vivo*, however, EpiSCs fall short (Brons et al., 2007). The most stringent way in mouse to assess pluripotency is the tetraploid complementation assay. A two cell embryo is fused to generate a tetraploid embryo that will continue to develop and can implant but will not form an embryo. When complemented with pluripotent cells one can create an "all ESC derived" mouse (Kang et al., 2009). A less stringent test of pluripotency is to inject ESCs into regular diploid blastocysts to form a chimeric mouse (Kang et al., 2009). Contribution to the mice as well as their ability to populate the germline is a measure of pluripotency. EpiSCs rarely contributed to chimeras and no germ line transmission was observed (Brons et al., 2007), which has been partly attributed to their different states and an inability to incorporate into the ICM rather than true differences in pluripotency.

From PGC to EGC

Reproduction is necessary for the survival of all species. In mammals genetic information is passed on to the next generation through the germline. In mice the majority of epiblast cells will continue on the path of differentiation, forming all the somatic tissues of the animal. A select group of cells, however, are set aside to become the precursors to the germ cells (Hayashi et al., 2007). Unlike other organisms, such as drosophila and C. Elegans, that set aside cells from the totipotent zygote to form the germ cells, the mouse germline is created from cells that are ostensibly on the path to a somatic fate. It is signals from the extra-embryonic ectoderm (ExE) and the visceral endoderm (VE) that provide the instructions for a small sub-set of epiblast cells to become primordial germ cells (PGCs) at 6.25dpc (Hayashi et al., 2007). As with ESCs it is BMP signaling which plays a large role in PGC specification. The ExE produces BMP4 and BMP8b, while the VE releases BMP2 (de Sousa, et al., 2004; Yoshimizu et al., 2001). These BMP signals result in the translocation of Smad proteins to the nucleus and the subsequent expression of fragilis, a nascent PGC marker (Saitou et al., 2002). Between 6.25-6.5dpc the expression of the key regulator of PGC specification, B-lymphocyte-induced maturation protein-1 (Blimp1), is detected in a small population of fragilis positive cells (Ohinata et al., 2005). Blimp1 functions not only to suppress the somatic program of these cells, but also to create an environment permissive to germ cell formation (Ohinata et al., 2005). In addition to Blimp1, the aforementioned pluripotent factors Nanog and Sox2 are re-expressed in PGCs as well as the PGC specific markers Stella and Nanos3 (Hayashi et al., 2007; Saga, 2008). As stated previously, PGCs are derived from a subset of epiblast cells that were on the path to become somatic tissue. Because of this, PGCs need to undergo an extensive reprogramming of their genome before they can assume a pluripotent state, including the reactivation of the Xi (de Sousa et al., 2008). This is achieved by extensive epigenetic remodeling which will be covered in more detail later. Suffice it to say that PGCs need to wipe the slate clean not just in relation to their somatic transcriptional program, but also in relation to their sex specific program, i.e. genomic imprints.

Genomic imprinting refers to the process by which genes are expressed or repressed depending on their parental origin. This means that some genes are active solely on the maternal allele and some expressed only on the paternal allele (Cattanach, 1986). Imprinted genes tend to play major roles in prenatal growth, placental development, bone and muscle development, and even post natal behavior (Mann et al., 1990; Coan et al., 2005; Plagge et al., 2004). It is not surprising then that misregulation of imprinted genes often results in growth abnormalities (Mann et al., 1990). When forming haploid gametes it is therefore necessary for mammals to erase the imprints inherited from their parents and reset them in their own sex specific fashion.

Much like the ICM and epiblast, PGCs can be explanted and under certain culture conditions give rise to a third pluripotent cell type embryonic germ cells (EGCs). Of these three pluripotent cell types, however, EGCs differ the most from their *in vivo* counterpart. While PGCs are lineage restricted and can only give rise to an organism's germ cells, EGCs exhibit all the *in vitro* pluripotent properties of ESCs and EpiSCs. In fact, when explanted, reprogramming from unipotent PGC to pluripotent EGC takes place over the course of approximately ten days (Durcova-Hills et al, 2008). Unlike EpiSCs, EGCs can be cultured and maintained in media conditioned with serum and LIF as well as in 2i, similar to ESCs (Leitch et al., 2010). However, in serum conditions it has been suggested to supplement the media with FGF-2 for the first 24 hours of derivation (Durcova-Hills et al., 2006). Blimp1, as mentioned previously, is responsible for the repression of the somatic transcriptional program in PGC development. In addition to suppression of somatic genes, Blimp1 functions to prevent a full pluripotent phenotype in PGCs. While PGCs express many pluripotent markers, i.e. Sox2, Nanog, Oct4, and Stella, two key pluripotent factors, C-Myc and Klf4, remain repressed (Durcova-Hills et al., 2008). Blimp1 associates with Prmt5, a histone methyltransferase, to repress expression of these genes in PGCs. Staining of PGCs in culture shows Blimp1 and Prmt5 co-localizing in the nuclei of PGC cells on day one. On day two , however, Blimp1 is no longer detected in the nucleus, and by day seven Prmt5 has been translocated to the cytoplasm. By day three, Klf4 is re-expressed and by day seven C-Myc is actively transcribed both of which are specific Blimp1 targets. It is believed that FGF-2 signaling in the first 24 hours of culture is essential for the repression of Blimp1, as its absence results in its persistent expression and failure of PGCs to reprogram into EGCs (Durcova-Hills et al., 2008). Blimp1 has been shown to co-localize with HDACs, indicating a dual repression of Blimp1 targets through both histone methylation and deacetylation (Yu et al., 2000). While EpiSCs showed limited contribution to chimeric mice, EGCs much like ESCs, contribute to all three germ layers as well as exhibiting germ line transmission (Labosky et al., 1994). While these chimeric mice were viable, their survival rate was directly correlated with the overall contribution of the EGCs. Mice with a small percentage of EGCs contribution survived. However, mice showing a larger amount of EGC distribution of 25-50% throughout their tissue tended to die prenatally at time point 14.5dpc, with those that do survive until birth dying within days (Tada et al., 1998). This lethality is attributed to the misregulation of imprinted genes in the EGCs, with many of the deceased offspring showing overgrowth.

Not many scientific fields are as susceptible to the political climate as the study of hESCs. From 2001 to 2008, federal funding for hESC research was severely reduced. Moreover, derivation of new hESCs lines after August 2001 was forbidden (Nature editorial, 2010). Mouse ESCs, EpiSCs, and EGCs (or mouse pluripotent stem cells (mPSCs)) not only give us an opportunity to study the mechanisms of pluripotency and development in depth, they also provide us with incredible insight into human embryonic stem cells (hESCs). Oct4, Nanog, and Sox2 are all required for the global transcriptional network in both mPSCs and hESCs (Babaie et al., 2007). Furthermore, Oct4 and Sox2 regulate their own expression in a cooperative fashion in both mouse and human as well as controlling the expression of Nanog (Okumura-Nakanishi et al., 2005; Koruda et al., 2005). Bizarrely, while Nanog is necessary for the establishment of pluripotency in both mPSCs and hESCs, it is not required for the self-renewal capabilities of mPSCs in culture but is indispensible in hESCs (Zaehres et al., 2005; Chambers et al., 2007). Other subtle differences between the mESC and mEpiSC cell types inform us about the nature of hESCs. Comparison of Oct4 target sites in mESCs and hESCs shows that there is a large difference in the regulatory circuits of these two cells types. However, when comparing the Oct4 target sites of mEpiSCs and hESCs there is approximately a seven-fold increase of common binding sites than in mESCs (Tesar et al., 2007). Culture conditions for hESCs also resemble EpiSC conditions more closely. While BMPs are an essential component for the maintenance of mESC pluripotency, in hESCs the addition of

BMPs encourages differentiation (Xu et al, 2002). In addition hESCs are completely unresponsive to LIF (Humphrey et al., 2004). However, much like EpiSCs, Activin/Nodal signaling plays a major role in the maintenance of hESCs, with the addition of bFGF proving essential for maintaining pluripotency (Wang et al., 2005). In addition, the same Xi seen in the mEpiSCs is exhibited in hESCs, although some variability in Xi is seen in hESCs depending on passage number (Dvash et al., 2010). There are undeniable similarities between mouse and human ESCs with regards to the main pluripotency network. Furthermore the more in-depth similarities between mEpiSCs and hESCs work only to reaffirm mESCs as a perfect model system for understanding the dynamics of human pluripotency.

Epigenetic Mechanisms

"Epigenetics" was first used by Conrad Hal Waddington in 1942 to help explain the ability of complex multi-cellular organisms to arise from a totipotent zygote via a process of cellular differentiation (Waddington, 1942). More recently epigenetics has been more precisely defined by C. Wu as "the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence'' (Wu and Morris, 2001) or as it is further described as a "stable yet *reversible* molecular mechanisms that lead to a given phenotype without a change in genotype (Meissner, 2010). This reversible quality or epigenetic marks is an important distinction and will be discussed further later in this thesis.

DNA does not exist solely as a double stranded helix. It must be packaged and regulated. This "packaging" comes in the form of chromatin, which is the combination of DNA and proteins. In the nucleus of eukaryotic cells DNA is wrapped around proteins called histones. These consist of an octomer of highly conserved proteins put together in pairs of the core histones H2A, H2B, H3, and H4. Approximately 147 base pairs of DNA are able to wrap around this octomer to form a nucleosome (Luger et al., 1997; Strahl and Allis, 2000). These nucleosomes are arranged into higher-order structures with the addition of the linker histone H1, in order to form chromatin (Bartke et al., 2010). The nucleosome is the context by which factors can regulate and modify chromatin in order to establish a cell's epigenetic identity. In addition to extrinsic and intrinsic signals activating and inhibiting various pathways, a cells identity is both maintained and lost through this epigenetic regulation. It is therefore necessary to take these epigenetic marks into consideration in order to fully understand the nature of pluripotency as well as mammalian development.

There are two main modes of epigenetic regulation that will be discussed here. One is to directly modify the DNA and the other is to modify the histones, around which the DNA is wrapped (Bird, 2002; Bernstein et al., 2006).

DNA methylation is one of the best-understood epigenetic mechanisms and found in all vertebrates as well as plants and other organisms (Feng et al., 2010). Methylation of cytosine is achieved by covalently attaching a methyl group to the 5-carbon position that is usually guanine nucleotide, otherwise known as a CpG dinucleotides (Jaenisch and Bird, 2003). In most mammalian cells, 70-80% of CpGs are methylated. Most methylation is found at repetitive elements such as transponsons, LINEs, SINEs, and centromeric heterochroamitn. There is another group of CpGs found 5' of the first exon of ~60% of human genes referred to as CpG islands (CGIs) (Bird, 1986). For the most part these CGIs remain unmethylated, and are instead regulated by repressive histone marks (Brinkman et al., 2012). Interestingly, CpG dinucleotides occur in the overall genome at approximately 1/5 of the expected frequency (Bird, 1986). The significance of these unmethylated CpGs remains unclear as many promoters can remain unmethylated regardless of transcription (Guenther et al., 2007). It has been shown that a protein named Cfp1 binds specifically to these unmodified CGIs *in vitro* and is shown to interact with the histone methyltransferase Setd1 to allow active marks to be deposited, and in turn create a chromatin environment more permissive to transcription (Voo et al., 2000; Thomson et al., 2010).

Deposition and maintenance of DNA methylation is carried out by three enzymes referred to as DNA methyltransferases (Dnmts). The first, Dnmt1, is tasked with propagating DNA methylation patterns faithfully after cell division. Dnmt1 performs its task through recognition of hemi-methylated CpGs and in turn deposits a methyl group on the cytosine of the nascent strand (Yang et al., 2010). Dnmt3a and Dnmt3b, are tasked with *de novo* methylation and act in conjunction with histone modifiers to silence genes (Karimi et al., 2011). In addition to promoter silencing, Dnmt3a/b help establish intragenic methylation, which as been shown to improve elongation efficiency as well as helps prevent initiating transcription from alternative promoters within the gene body (Lorincz et al., 2004; Maunakea et al., 2010). A homologue of Dnmt3a/b, Dnmt3l has been shown to be necessary in the establishment of DNA methylation at imprinted genes during oogenesis (Bourc'his et al., 2001). Methylation patterns in somatic cells remain mostly unchanged, with only isolated fluctuations dependent on the activation or repression of specific cellular processes. There are instances when global demethylation are the norm. This can be seen in development during two distinct phases, more specifically pre-implantation development as well as germ line formation (Santos et al., 2002; Seki et al., 2005).

PGCs undergo dramatic global demethylation of their genome and most promoters, intergenic regions, and transposon sequences are demethylated by the end of their reprogramming at 13.5dpc. The only exception to global erasure is intracisternal A particles (IAPs), a family of retrotransposons that exhibit a unique resistance to demethylation (Popp et al., 2010). This global demethylation includes the removal of the aforementioned imprinted genes. Imprinted genes are identifiable by their differentially methylated regions (DMRs). For example, for the imprinted gene H19 the paternal allele is methylated and remains silent while the maternal allele is unmethylated and expressed (Bartolomei et al., 1991). The DMR of H19 is located ~4kb upstream of the H19 promoter. In imprinted genes the DMR is referred to as the imprint control region (ICR) since its deletion results in the misregulation of the imprinted gene (Thorvaldsen et al., 1998). Methylation at ICRs can act to either recruit factors that can only bind methylated cytosines, such as MeCP1 and MeCP2, or that can recognize methylated cytosines and bind them, thereby repressing transcription (Jones et al., 1998). Alternatively methylation may act to block the access of proteins to the DNA. In H19 the methylation of the ICR acts to inhibit the insulator protein CTCF from binding. CTCF binding on the unmethylated maternal allele allows CTCF to create a boundary between the downstream enhancers of the Igf2 maternal allele, blocking its transcription while allowing for the expression of the H19 non-coding RNA (Mancini-DeNardo et al., 2003; Hark et al, 2000). Only after all the ICRs are removed can the PGCs then reassemble their methylation patterns in a sex specific manner. How these imprints are reestablished in such a tightly regulated manner is still unclear, however certain factors have been identified. While it has been shown clearly that Dnmt3a and Dnmt3l are required for the methylation of ICRs, how they are directed to these regions is still somewhat unclear (Bourc'his et al., 2001). Some evidence suggests that the zinc finger protein ZFP57 is required for the establishment of ICRs such as Snrpn. ZFP57 has also been shown to be dispensable for the establishment of others (Li and Sasaki, 2011). What is clear is that failure to reestablish DNA methylation patterns in germ cells will result in embryonic lethality of subsequent offspring. In male gamete production a lack of methylation results in serious defects in spermatogenesis as well as sterility. In female gametes, however, fertility is not affected, with that in mind, embryonic lethality is still unavoidable due to the defects in imprinted genes (Hata et al., 2002; Kaneda et al., 2004). These differing phenotypes illustrate differences in the two sexes with regards to their tolerance of a hypomethylated genome. This intriguing phenotype is indicative of the methylation dynamics witnessed in normal mammalian development.

During development the DNA methylation profiles of the sperm and the oocyte are remarkably different (Smallwood et al., 2011). While the sperm is hypermethylated the oocyte itself is incredibly hypomethylated, this hypomethylation even extends to repetitive element such as LINEs, SINEs, and long terminal repeats. As with PGCs the IAPs in oocytes are also more methylated (Kobayashi et al., 2012). These drastic differences in the methylomes of the sperm and the oocyte are soon reconciled after fertilization when the sperm genome undergoes drastic demethylation (Oswald et al., 2000). This includes the demethylation of many of the retroelements, such as LINEs and LTRs (Kim et al., 2004). Why these elements become demethylated, thereby risking reactivation, is not entirely clear, but there is a great deal of speculation that reactivation of retroelements may plays a role in evolution as many of these elements are able to regulate gene expression (Shapiro, 2010; Shapiro, 1969). What is clear from this rapid demethylation of the sperm DNA is that there something present in the oocyte which is capable of wide spread demethylation.

For many years the mechanism by which DNA was demethylated was a hotly debated topic. There were two schools of thought when it came to this question. The first camp believed that the process was an active one, that is, a hitherto unidentified demethylase was either removing the methyl group from the cytosine or that the entire cytosine base was being removed and replaced with an unmodified cytosine (Wu and Zhang, 2010). The second camp argued that the process was completely passive. They argued that methylated cytosine was simply diluted out over multiple cell divisions (Kagiwada et al., 2012). As with most debates the answer lay somewhere in between. For the most part DNA demethylation, especially global demethylation, is a passive process. This was detailed by Mitinori Saitou in 2012, what this research showed was that while the maintenance enzyme

Dnmt1 was present in PGCs, its' cofactor Uhrf1 was absent resulting in Dnmt1 being unable to recognize hemimethylated DNA after each cell division. In addition to this the double time of PGCs was much shorter than previously thought, allowing for a more rapid dilution of methylated cytosine (Kagiwada et al., 2012). While not exactly active, a form of targeted demethylation exits. The family of enzymes known as Ten-Eleven Translocation enzymes (Tet) are able to oxidize 5-methylcytosine into 5-hydoxymethylcytosine (Tahiliani et al. 2009). There are three different Tet enzymes, each performing the same task but at different times. Tet2 is required for efficient reprogramming of somatic nuclei when fused with pluripotent cells. Tet1 has been identified as a necessary component of imprinted gene demethylation, and Tet3 is required for the rapid demethylation of the sperm DNA upon fertilization (Piccolo et al., 2013; Iqbal et al., 2011). In fact it has been shown that the sperm pro-nuclei exhibits a global conversion of 5-methylcytosine to 5- hydroxymethylcytosine and that this mark is completely absent from the female pro-nuclei which is protected from oxidation by Dppa3 (Gu et al., 2011). Rapid demethylation of the sperm upon fertilization occurs as Dnmt1 is unable to recognize this mark and it becomes diluted over subsequent cell divisions (Iqbal et al., 2011).

While there are only four main types of DNA modifications, histone marks come in a wide variety and can help to delineate different functional elements of the genome (Mikkelsen et al., 2007) Transcription is a tightly regulated process, but often times its' complexity is over looked. The central Dogma of all biology simply states that DNA is transcribed into RNA which is translated into protein, (ncRNA not withstanding) (Mattick, 2003). Upon closer inspection things become more complicated. A genes journey to functional protein relies on many different elements: promoters, gene bodies (both exons and introns), enhancers and boundary elements all play roles in this process and all must be differentially regulated and this is carried out by differential modification of the histones contained in the nucleosome. Most important for a genes transcription is the promoter, if the promoter is not accessible to transcription factors (TFs), then no transcription will be initiated.

H3 lysine 4 trimethylation (H3K4me3) is the mark most associated with the transcription start site (TSS) of many actively transcribed genes (Henikoff et al., 2004). The previously mentioned Cfp1, that has been shown to bind unmethylated CpGs, has also been shown to recruit the MLL family of histone methyltransferases required for deposition of H3K4me3. Knock out of Cfp1 results in an absence of this active mark from the promoters of genes (Clouaire et al., 2012). Additionally the presence of H3K4me3 at promoters is sufficient to inhibit the de-novo Dnmt3 family of methyltransferase via steric hindrance of their binding domain (Otani et al., 2009). In addition to H3K4me3, acetylation of histone 3 (H3) has been shown to co-localize with H3K4me3 at the promoters of expressed genes, indicating its' role as a mark of active transcription, H3 acetylation was also found in intergenic regions of active genes as well (Roh et al., 2005). This also suggests a certain level of redundancy in the cells maintenance of gene expression (Schubeler et al., 2004). While active marks are necessary for gene expression there must be a mechanism for the repression of genes, DNA methylation is mostly considered to be a stable repressor of transcription, however, there are instances when it is absent from the promoters of genes that are inactive regardless. Histones can be modified in such a manner as to make transcription impossible, the most well known repressive modification is H3K27me3 (Mikkelsen et al., 2007). This mark is deposited by the Polycomb group proteins (PcGs), Polycomb group (PcG) proteins were initially discovered in Drosophila melanogaster as being required to prevent errant expression of the developmental homeotic (Hox) genes (Lewis, 1978). Biochemical studies show that PcG proteins assemble and function in multi-protein complexes. PcG silencing involves at least two kinds of multiprotein complexes that work together. These are known as the PRC1 and PRC2 complexes. A reconstituted complex of the proteins in the PRC1 complex can alter chromatin and DNA structure, inhibit chromatin remodeling, and inhibit transcription (Francis et al., 2001). Components of PRC1 can also function as an E3 ligase for histone H2A ubiquitylation, all of these activities have been shown to correlate with silencing (Cao et al., 2005). The PRC2 complex contains the components necessary for the deposition of the repressive mark H3K27me3, this mark is then recognized by components of the PRC1 which then results in the stable silencing of transcription (Wang et al., 2004). Just as DNA methylation patterns are maintained through cell divisions via the maintenance methytransferase DNMT1, PcG proteins have been shown to bind to chromatin during the S-phase of replication as well as the newly synthesized DNA, highlight the importance of PcG proteins in epigenetic inheritance (Francis et al., 2009). Thusly PcGs are essential for the maintaining tissue specific repression of non-essential genes (Simon et al., 2009).

Overall it is clear that histone modifications and DNA methylation are important for the over all fidelity of a cells transcriptional program and for development overall. ESCs exhibit unique epigenetic qualities, while DNA methylation in somatic cells is deposited only in a CpG context, ESCs also exhibit limited non-CpG methylation, a phenomenon also witnessed at the early stages of development (Ramsahoye et al., 2000). The reason for this non-CpG methylation is not fully understood and it is mostly reconciled upon differentiation (Ramsahoye et al., 2000). ESCs also provide insight into the connectivity and interplay between the key pluripotent factors and epigenetic remodeling complexes, and how these maintain a delicate balance between pluripotency and multi-lineage differentiation. Oct4 has been shown to play important roles in the 3D structure of chromatin via interaction with CTCF. These interactions may function to bring the genes that encode NANOG and developmental pluripotency-associated protein 3 (DPPA3) into an active transcription zones via chromatin looping. Loss of Oct4 results in the collapse of these 3D structures (Levasseur et al., 2008). ESCs also show us that histone marks themselves are not sufficient for the formation of heterochromatin but instead serve to provide a target for other repressive complexes. Recently it has been shown that levels of the repressive histone mark H3K9me2 are similar between ES cells and differentiated cells in mice. H3K9me2 becomes more punctate after differentiation of mouse ESCs. This data suggests that heterochromatin reorganization during differentiation is not due to an increase in the repressive histone mark but instead is a result of the reorganization of the mark, and that other factors, as well as histone modifications, must play a role in the remodeling of heterochromatin in ESCs (Lienert et al., 2011). ESCs exhibit unique feature at the promoters of approximately twenty percent of promoters of genes are associated with PRC2 and contain the repressive H3K27me3 histone mark (Mikkelsen et al., 2007). Surprisingly these same promoters also exhibit the active histone mark H3K4me3 (Ku et al., 2008). These promoters are referred to as “bivalent.” These bivalent genes are also sometimes seen as being “poised” as a large number of these genes are associated with early developmental pathways, causing many to speculate that this bivalency may exist to allow easy activation or suppression of these genes, as evidenced by many of these genes having their bivalent state resolved at key stages corresponding with their developmental activation (Mikkelsen et al., 2007). In addition to the resolution of bivalent domains, the histone and DNA methylation landscape changes drastically upon differentiation (Gifford et al., 2013). While it is necessary to shut down the many factors associated with pluripotency, many changes in DNA methylation occur upstream from most promoters, possibly indicating that DNA methylation is responsible for modulating enhancers and other regulatory elements (Meissner et al., 2008). As tightly regulated as epigenetic mechanisms are, it is surprising to discover that the self-renewal of stem cells is unaffected even under the most drastic misregulation. Knock-outs of the DNA methyltransferases as well as essential components of the Prc2 complex resulted in no loss of self-renewal in ESCs, however differentiation capabilities were lost *in vitro* as well with embryonic lethality being the prevalent phenotype *in vivo* (Jackson et al., 2004; Tsumura et al., 2006; Chamberlain et al., 2008). Differentiation is with no doubt dependant on global and coordinated epigenetic processes the same is true for the reverse i.e. reverting a terminally differentiated cell back towards a pluripotent state.

Reprogramming

Before the field of epigenetics was established it was unclear what was occurring during the development from zygote to multi cellular organism. It was speculated that DNA was parsed out in a cell specific manner with each cell type receiving only the genes necessary for it function. That was until 1953 when Briggs and King, using the frog species Rana-Pipiens, took the nucleus of an embryonic cell and placed it into the enucleated zygote (Briggs, 1953). Normal development was surveyed in approximately 35% of zygotes with a nucleus transplanted from the blastula, however this efficiency dropped drastically when the nuclei of later developmental cells were used i.e. the late gastrula endoderm. Cells from later stages failed to develop at all. On the surface these experiments may lend credence to the notion that the genome of all cell types are not equipotent. In 1958 Gurdon was able to repeat these experiments and successfully produce live frogs from later developmental stages (Gurdon et al., 1958). It was still unclear as to whether or not these cells were in fact fully differentiated, however in 1975 Brun et al. were able to produce live off spring using the nuclei of fully differentiated lymphocytes (Brun and du Pasquier, 1975). In mammals the first report of a successful nuclear transfer to an enucleated zygote was in 1981 by Hoppe et al. they claimed that transfer of the nucleus from an ICM cell of a mouse was successful in developing to live birth (Ilmensee and Hoppe, 1981). Unfortunately these results were not reproducible and it wasn't until 1986 that Steen Willadsen became the first person to officially clone a mammal (Willadsen, 1986). Where Willadsen succeeded, when others had failed was down to his use of an oocyte that had been arrested at the metaphase stage of the second meiotic division, additionally the spindle and metaphase plate had been removed causing speculation that exposure to oocyte cytoplasm was necessary for success, perhaps due the donor nucleus needing time to adjust to the environment or for factors present in the oocyte to act on the donated nucleus. This theory gained more credence as it became clear that successful cloning was greatly helped by performing serial transplantation (Tsunoda and Kato, 1998). Serial transplantation is the technique by which donor nuclei are initially fused with enucleated oocytes, after being developed to the two-cell stage, after which they are then enucleated and the removed nuclei are fused with yet another enucleated normal two-cell-stage cytoplast. In addition to serial transplant, efficiency of nuclear transfer was greatly increased by selecting donor nuclei of cells in the G1 phase of the cell cycle (Campbell at al., 1993). In normal development the sperm and egg genome are transcriptionally silent upon fertilization. It is therefore necessary their chromatin to be extensively remodeled in order for the developmental program to be activated. Nuclear transfer is no different, the donated nuclei must be reprogrammed in order to active it's developmental transcriptional program, however unlike sperm and egg nuclei, the donor has the additional hurdle of the already being transcriptional active and the oocyte cytoplasm need to carry out extensive nucleosomal reprogramming (Kikyo et al., 2000). In addition to histone modifications, DNA methylation status of the donor embryo plays a large role in reprogramming. The increased success rate of the early nuclear transfer experiments relied heavily on taking a nucleus from an early stage of developments, this is important as key pluripotent genes will still be active at this stage and free of DNA methylation (Simonsson and Gurdon, 2004; Hattori et al., 2004). It has been shown that the methylation status of the donor cell plays a large role in the success of nuclear transfer reprogramming. Cell lines containing a mutated version of Dnmt1, resulting in a globally hypomethylated genome, more readily gave rise to three times the number of successful clones in nuclear transfer experiments. Additionally nuclear transfer of nuclei from neural stem cells that already express Oct4 and Nanog improved efficiency as well (Blelloch et al., 2006). All of these groups, while successful, were still transferring the nuclei of cell at an early stage of development. It wasn't until the cloning of Dolly the sheep that it became clear that not just embryonic nuclei but adult fully differentiated nuclei could be fully reprogrammed towards totipotency (Wilmut et al., 1997).

The nuclear transfer experiments show that factors present in the oocyte have the ability to impose their epigenetic profile on the introduced somatic nucleus. Similar results can be witnessed when pluripotent cells are fused with somatic cells. As preciously discussed PGCs, undergo a genome wide epigenetic overhaul *in vivo*, and the *in vitro* EGCs derived from PGCs exhibit similarities in their epigenotype as evidenced by their lack of DNA methylation at ICRs. In 1997 Surani demonstrated that by fusing EGCs with somatic lymphocyte cells the EGCs was able to impose it's epigenetic profile on the somatic DNA, this included loss of DNA methylation at numerous loci as well as the loss of methylation at imprinted genes

(Tada et al., 1997). These cells maintained their pluripotent morphology while in culture and upon blastocyst injection these tetraploid cells were actually able to contribute to multiple tissue types. Tetraploid complementation assays, however, did not result in the birth of live pups due to the obvious misregulation of imprinted genes, and the embryos exhibited extreme morphological abnormalities (Tada et al., 1997). In 2001 Tada fused ESCs with somatic thymocyte cells, this time with more of a focus on the actual acquiring of pluripotency in the hybrid cells as opposed to sheer epigenetic remodeling. As stated before, in female cells an X chromosome is selected for random inactivation and as mentioned, mESCs exhibit two Xa, while PGCs reactivate the Xi of the epiblast cells. When ESCs are fused with female somatic cells the Xi once again becomes active as evidenced by Xist expression, an active X shows a small spot of Xist expression whereas an inactive X will be completely coated (Tada et al., 2001). This is reminiscent of the reactivation of the X chromosome in nuclear transfer experiments (Egan et al., 2000). In addition to X reactivation, an Oct4 transgene in the somatic genome became active within 48 hours of fusion and much like the EGC/somatic fusion, contribution to all three germ layers was witnessed *in vitro* in chimeras. Unlike EGCs, ESCs maintain their differentially methylated ICRs and therefore the imprints of the somatic cells remained unchanged. Indication that this is a feature unique to EGCs and a remnant of their *in vivo* PGC program. Additional fusion experiments have provided more evidence for the role of pluripotent factors in genome wide reprogramming. Smith et al. have shown that over expression of Nanog can help increase reprogramming efficiency, however it is not sufficient in of its' own. Over-expressing Nanog alone was not enough to induce a pluripotent state in neural stem cells, however when these nanog expressing cells were fused with ESCs the reprogramming efficiency shot up. Additionally the over expressing of nanog in the ESCs was enough to increase the formation of pluripotent hybrids by ~ two fold (Silva et al., 2006). Conversely the loss of Oct4 expression in ESCs results in the lack of reprogramming in heterokaryons (Pereira et al., 2008).

While the fusion of a pluripotent cell with a somatic cell, can offer some insight into the reprogramming process, the tetraploid nature of these cells is undesirable. It was therefore necessary to develop a process to reprogram somatic cells in such a way that they could be biologically useful. After all it is the main goal of the stem cell field to help further regenerative medicine. Not only to have access to multiple cell types in unlimited quantities, but to be able to create these cells in a patient specific manner. With the availability of (hESCs) and improved protocols for their differentiation, this goal is becoming realized more and more with each passing year. There have been great leaps forward in the field and these indicate that the over all goal of directed differentiation could indeed translate into effective treatments for many human diseases (Shiba et al., 2012; Wang et al., 2013; Ma et al., 2012). The ethical issues surround the derivation of hESCs, mainly concerning the destruction of human embryos, have stymied much of the research in this field, however the main obstacle is the creation of patient specific i.e. isogenic cell lines. A huge problem with organ transplantation is the host's rejection of the donor's organ, this is what the regenerative field seeks to resolve. The creation of isogenic organs in the laboratory, instead of allogenic donors could help resolve the rejection issue

(Flanagan et al., 2010). In 2006 Yamanaka addressed this question with the creation of the first induced pluripotent stem cells (iPSCs). In the late 1980s Gehring et al. were the first to show that the developments fate of a cell could be altered by the ectopic expression of genes. In their experiments they showed that over-expression of the gene *Antennapedia* in *Drosophila* could induce the growth of legs in place of antenna (Schneuwly et al., 1987). Again in the mid nineties Gehring emphasized this point by once again over-expressing another gene, this time *eyeless,* t resulting in the development of wyes on the legs and wings of the hapless *Drosophila* (Gehring, 1996). In mammals Weintraub showed that ectopic expression of MyoD in mouse fibroblasts was sufficient to push these cells towards a myoblast fate (Davis et al., 1987). Using these studies and the results from the numerous fusion experiments, Yamanaka hypothesized that the same techniques could be utilized to convert a somatic cell to a pluripotent state. Twenty-four candidates were selected based on their roles in maintaining pluripotency, including the dominant negative form of Grb2, a gene known to negatively affect pluripotency (Burdon et al., 1999). Inserting a resistance gene into the pluripotency related gene Fbx15 that would only become active upon acquirement of pluripotency, the twenty four factors were introduced retrovirally into MEFs. With all 24 factors pluripotent colonies were produced. Stepwise removal of certain factors identified four sufficient to induce pluripotency. These four factors were Oct4, Sox2, Klf4 and c-myc. Removal of Oct4 and Klf4 resulted in no pluripotent colonies. Removal of Sox2 resulted in reduce efficiency of reprogramming as evidenced by fewer pluripotent colonies forming in its' absence, while withdrawal of c-myc resulted in colonies, these colonies where flatter in morphology. All four of these factors proved necessary for iPSC formation. Oct4, Klf4, and c-myc with the absence of Sox2, resulted in colony formation, however these were morphologically different from the iPSCs derived with all 4 factors, implying that while the Fbx15 gene was activated full reprogramming had not taken place. Somewhat surprisingly, Nanog was not essential for reprogramming (Takahashi and Yamanaka, 2006). Despite the ESCs morphology of the different iPSC lines derived, not all exhibited the same level of pluripotency. Some cells could produce tumors *in vivo* exhibiting all three germ layers, while others could not. Blastocyst injection showed contribution to all tissues, germline transmission was unclear and no live chimeric pups were born. Micro array data also showed large difference in the gene expression of ESCs compared with iPSCs. While the true pluripotent nature of these iPSCs was questionable, what this work ultimately did was further dissect the mechanisms of pluripotency. The roles of c-myc and Klf4 were of particular interest. C-myc has been shown to interact with histone acetyltransferases raising the question of whether c-myc induced histone acetylation is responsible for the open chromatin state of many Oct4/Sox2 binding sites (McMahon et al., 1998; Fernandez et al., 2003). Additionally Klf-4 has been shown to inhibit p53, a known inhibitor of Nanog, perhaps this is why Nanog is dispensable for iPSC generation (Rowland et al., 2005). In 2007 Yamanaka applied the same four factor strategy used in mouse iPSCs to human reprogramming with the same success (Takahashi et al., 2007). After derivation of human induced pluripotent stem cells (hiPSCs) these cell were then successfully differentiated toward a cardiomyocyte fate. Once again microarray analysis showed that hiPSCs while being similar to hESCs are still not identical. Some rudimentary chromatin analysis showed the gain of H3K4me3 and the loss of H3K27me3 at the promoters of pluripotent genes, as well as the return of bivalent domains at the promoters of developmentally associated genes, reflective of hESCs (Pan et al., 2007). More in-depth genome wide analysis has shown extensive global H3K4me3 (Koche et al., 2011). As fascinating as the iPSC field is, the process itself is still incredibly inefficient. Many factors could contribute to this but it appears the biggest barriers to reprogramming re epigenetic. Transient knock down of Dnmt1 greatly improves reprogramming efficiency, highlighting the necessity for the demethylation of genes associated with pluripotency (Mikkelsen et al., 2008). The biggest draw back to inducing pluripotency in somatic cells is related to the integration of the retroviruses. Analysis of iPSC clones identified three to six retroviral integrations of each factor, this means up to twenty four random and possible mutagenic insertions leading to an increased chance of tumor formation. In fact it has been shown that mice derived form iPSCs had an increased chance of developing tumors and that this tumerogenesis could be traced back to the errant reactivation of the c-myc retrovirus (Okita et al., 2007). These dangers are not surprising when you consider the oncogenic properties of c-myc and Klf-4 (Dang, 2012; Yu et al., 2011)). Obviously this must be reconciled before any possibility of human trials can take place. Many alternative methods of reprogramming have arisen to try and address the problem of maintaining genome integrity; excisable vectors, non-integrating vectors, mRNA, and recombinant pluripotent proteins (Somers et al., 2010; Woltjen et al., 2009; Stadtfeld et al., 2008; Warren et al., 2010; Zhou et al., 2009).

Unfortunately all of these alternative techniques share one common trait and that is their abject inefficiency (Stadtfeld et al., 2010). More recently researchers are turning to small molecule based reprogramming. Of all the four factors Oct4 has always proven to the most difficult to replace with chemical compounds, however, recent work has detailed the reprogramming of somatic cells utilizing only small molecules (Hou et al., 2013). It seems like it's only a matter of time before efficient reliable reprogramming is achieved.

RESULTS AND DISCUSSION

Experimental Approach

Despite similar pluripotent properties EGCs and ESCs are patently different (Figure S1). Not only are they derived at drastically different developmental time points, EGCs have been reported to exhibit a vastly different methylation profile compared with ESCs, though these reports have largely centered on low resolution "global" analysis such as genomic digestion with methylation sensitive restriction enzymes and Southern blotting (Tada et al., 1998). To better assess the differences in methylation status of these two equivalently pluripotent but discretely derived cell types, we performed an in depth comparison of EGCs and ESCs. The first goal was to determine which factors were present in EGCs that could account for their comparatively hypomethylated genome as well as their ability to demethylate regions that are usually recalcitrant to demethylation, such as the ICRs of imprinted genes (Tada et al., 1998). It was hypothesized that a transcriptional comparison employing next generation sequencing technologies for both coding and non-coding transcripts would uncover differences between the two cell types hitherto undiscovered. This approach is not without merit, various groups have approached this problem in a similar fashion, and while certain differences in gene expression have been observed, no smoking gun has yet been reported (Sabour et al., 2011; Mise et al., 2008). It should be said that many of these groups utilized a somewhat outdated microarray approach to gauge differential gene expression, which lacks the sensitivity of RNA-seq as well as the inability to measure non-coding RNAs.

In addition to differential gene expression, we wanted a more detailed picture of the methylation profile (methylome) of the EGCs and ESCs. Earlier papers such as Tada et al. tended to focus heavily on the differential methylation of ICRs, with some limited data regarding the methylation status of minor satellite repeats. In this study, we wanted to utilize the power of reduced representation bisulfite sequencing (RRBS) to paint a more detailed picture of the EGC and ESC methylome on a genome "scale" and at single nucleotide resolution, including that of promoters and coding sequences, but also that of repetitive elements such as Long Terminal Repeats (LTRs), Long Intergenic Nuclear Elements (LINEs), and Short Intergenic Nuclear Elements (SINEs).

Results

Our first objective was to isolate EGCs and ECSs from mice with identical genetic backgrounds to reduce "noise " from genetic variations that may occur naturally across various strains of mice. To this end, we crossed C57BL/6 females with 129 males. We selected this cross because of its high ESC derivation efficiency, as well as the robustness of this genetic background with regards to chimera contribution in blastocyst injections. ESCs were derived (see methods) at 3.5dpc, while EGCs were derived at two time points, 11.5 and 12.5 dpc. Most of the data presented stems from the comparison of EGCs at 12.5 dpc, from which we derived 3 male lines, two female lines, and an XO female line that spontaneously jettisoned one of its X chromosomes, a phenomenon not uncommon in the derivation of female stem cells. For ESCs, three female and three male lines were derived. After acquiring these cell lines, total RNA and genomic DNA samples were prepared for both RNA-sequencing and RRBS analysis, respectively (Methods).

RRBS permits a genome "scale" assessment of CpG methylation at a consistently sampled subset of CpGs. Instead of sequencing the entire genome, which is costly and primarily samples CpG poor regions where DNA methylation is not dynamic, restriction digestion that targets a CpG containing motif followed by size selection for short fragments selectively enriches for CpG dense regions. After sequencing and alignment of libraries generated from all of our cell lines, we were able to compare, the methylomes of our two cell types at high resolution. Our primary results were unexpected: while we expected hierarchical clustering of the data to separate cells based upon their cell type (ESC vs. EGC), they instead clustered according to the sex of the cell line (Figure 1).

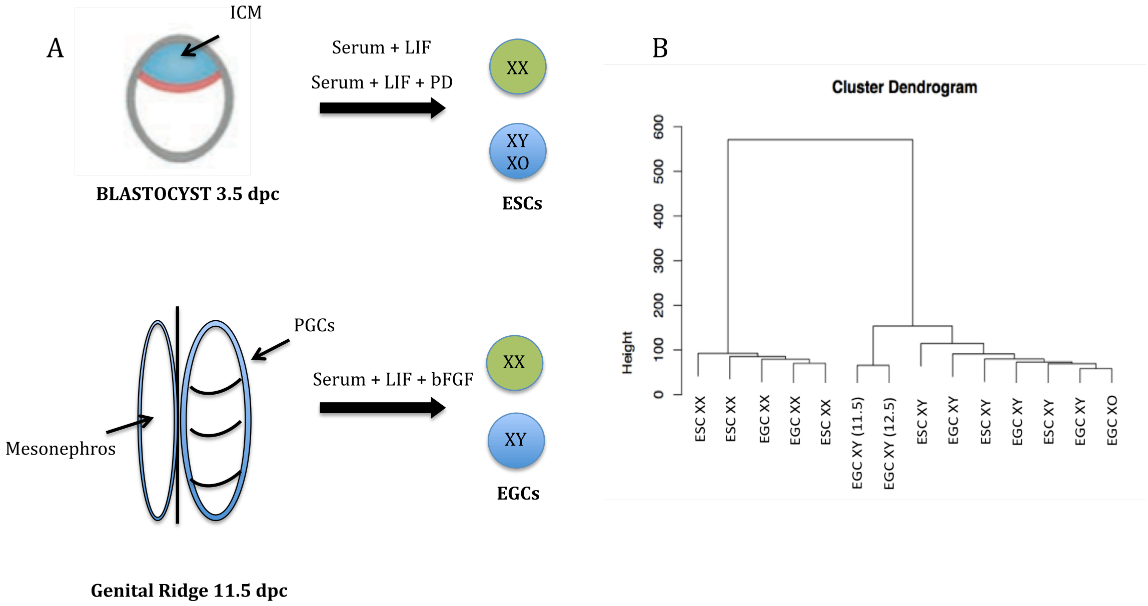


Figure 1. Hierarchical clustering of EGCs and ESCs. (A) Developmental origin of ESCs and EGCs and culture conditions. (B)Hierarchical clustering of the methylation profiles of EGCs and ESCs shows grouping of samples based on sex as opposed to cell type. The XO EGC sample can be seen to cluster with the male samples.

Analysis of 1 kilobase (KB) tiles demonstrates marked differences in the methylation profiles of male and female cell lines (Figure S2). We identified 27830 1KB tiles that were differentially methylated between the two cells by >20%, with the female cells showing substantial hypomethylation compared to the more hypermethylated male lines (Figure 2). This drastic difference in methylation profiles has been reported previously, though at lower resolution. In 2004, Brockdorff et al., perplexed by the frequent loss of the second X during the derivation of female ESCs, decided to investigate further differences between female and male that may explain this trend. Using rudimentary restriction enzyme and Southern blot analysis, they reported that female cells exhibited broad hypomethylation compared to males. With this in mind, we expected to see differences in the male and female ESC lines. We did not expect that this

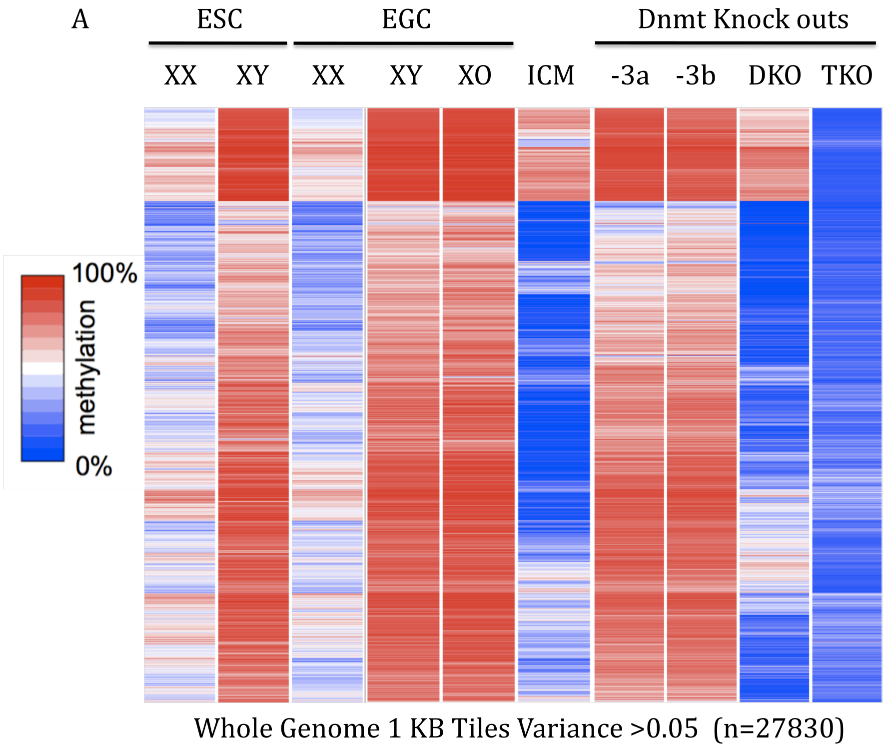


Figure 2. Genome scale 1 kilo base (Kb) tiles of ESCs and EGCs. Genome scale 1KB tiling of ESCs and EGCS shows 27830 tiles. Comparison with Dnmt knock out cells and ICM samples indicates an increased "naïve" pluripotency in female versus male pluripotent cell types. XO samples exhibit a methylation profile similar to that of the male cells.

trend would carry over to female EGCs, with stem cells derived from either PGCs or the ICM clustering by sex instead of origin. Previous claims about EGCs have always focused on their astounding lack of methylation, a phenotype that has always attributed to their derivation from PGCs. Our data, however, strongly suggest that these historical claims may have been misinterpreted. In addition to showing that EGCs do not represent a unique form of pluripotency as believed, we can also look in greater detail at the methylation profile of different genomic regions to see what discriminates male from female lines.

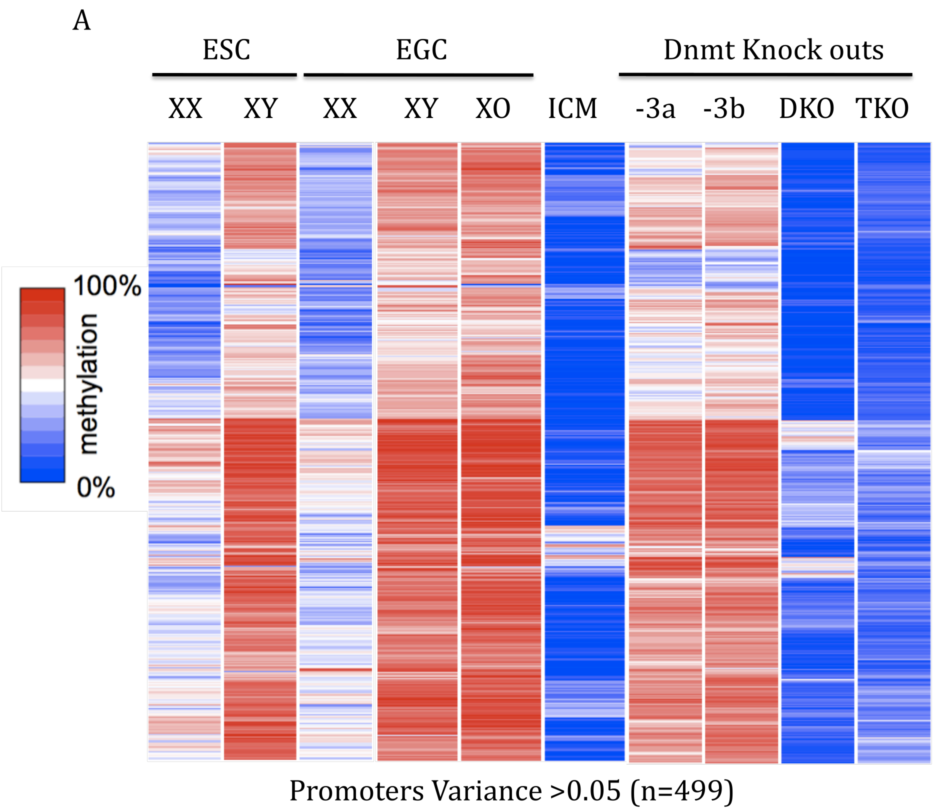


Figure 3. Promoter methylation profiles of EGCs and ESCs. (A) Comparison of the methylation at promoters shows that the majority of promoters in the cells remain unmethylated. A subset of promoters exhibit hypomethylation in female cells while remaining heavily methylated in male cells. XO samples once again exhibit methylation profile similar to that of male cells.

Promoters

The majority of CpGs in the genome are methylated and distributed intergenically or across gene bodies. Alternatively, CpGs found in promoters are mostly unmethylated, with only a small percentage that are dynamically methylated in a tissue specific manner (Weber et al., 2005; Ball et al., 2009). With this in mind, we decided to investigate and compare the promoter methylation between the two samples to search for an obvious regulatory answer to the difference in global methylation (Figure 3). Our data show that, while the majority of promoters in the male and female cells remain unmethylated, a subset of 499 promoters are specifically hypomethylated in female cell lines (Figure 3A)

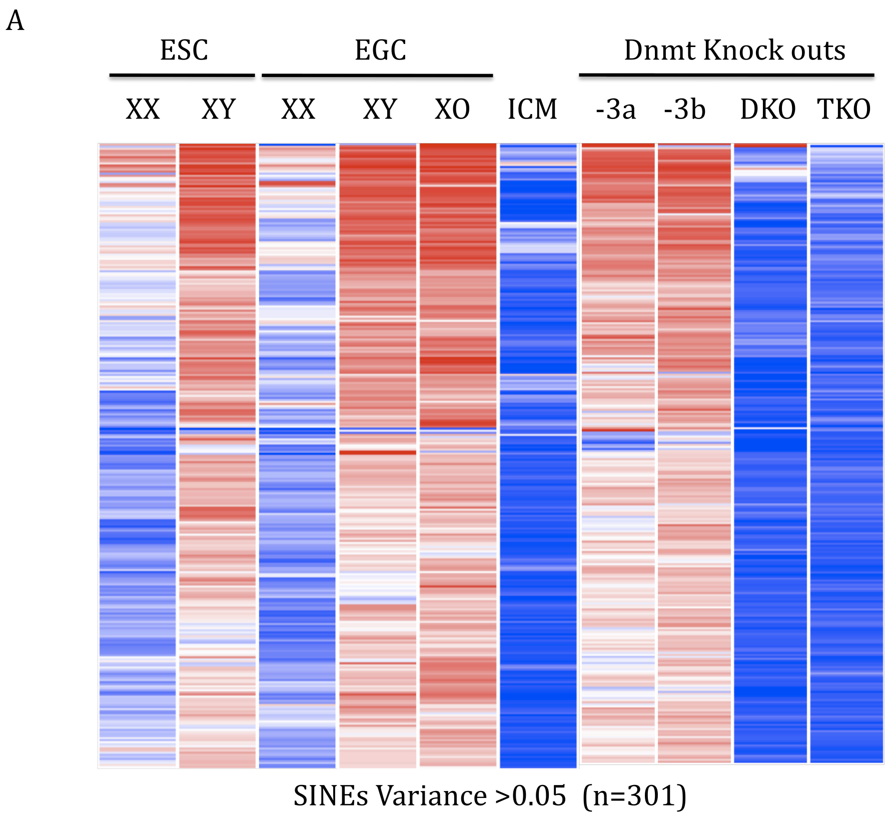


Figure 4. Methylation profile of SINEs in EGCs and ESCs showing variance >0.05. SINEs exhibit hypomethylation in female cells in contrast to males. Dnmt 3a/b double knockout cells and ICM samples exhibit extreme hypomethylation. XO samples resemble male samples.

SINEs LINEs and LTRs.

SINEs (short interspersed nuclear elements) are retrotransposable elements first discovered in 1981 (Jagadeeswaran et al., 1981). SINEs, along with other transposable elements, have long been considered "selfish" or "junk" DNA. More contemporary research is showing evidence of functional genomic roles for these elements in gene regulation and chromatin organization (Ichiyanagi et al. 2011). Considering that there are literally millions of SINEs, on average compromising

~10% of the mammalian genome, it is not surprising that they were so cavalierly dismissed (Elsik et al.,2009; Lander et al., 2001). When inserted into the introns of genes, SINEs play a large role in the alternative splicing of the messenger RNA (mRNA), creating unique protein products (Sorek et al.,2002). This is particularly evident in the human genome, where the SINE family *Alu* has been identified as a possible driver of human evolution with over half a million copies located in the introns of human genes (Gotea et al., 2006). In addition to alternative splicing, SINEs have been shown to function as distal enhancers, but more strikingly can remodel the chromatin landscape via the recruitment of nucleosomes (Bejerano et al., 2006). By affecting the nucleosome profile of specific genomic regions, either at enhancers or within the gene body, SINEs appear to function by impeding transcription (Tanka et al., 2010). Furthermore, SINEs can act as binding sites for chromatin boundary associated proteins such as CTCF. (Phillips et al., 2009). This creates a barrier between heterochromatin and euchromatin, helping to maintain the integrity of a cells epigenome at the nuclear level. A fundamental attribute of SINE-based gene based regulation may be its dependence on their methylation status, making them a germane focus for direct investigation within our samples. SINEs that contribute to the promoter sequences may affect gene repression by their methylation, and loss of methylation specific to these elements may result in spurious, deregulated transcription. This becomes important during the formation of the germline where, when compared to somatic cells, DNA methylation is

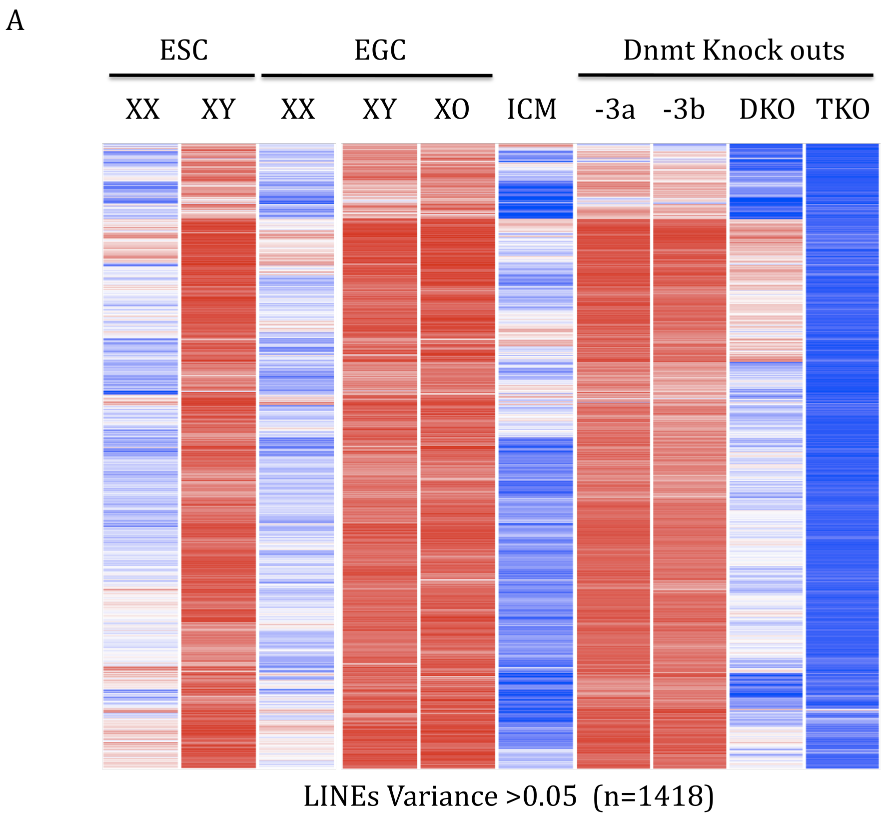


Figure 5. Methylation profile of LINEs in EGCs and ESCs showing variance >0.05 LINEs exhibit hypomethylation in female cells in contrast to males. Dnmt 3a/b knockout cells and ICM samples exhibit extreme hypomethylation. XO samples resemble male samples. Female cells more closely resemble Dnmt 3a/b double knockouts than the ICM.

incredibly low; this is also the case during preimplantation development (Ichiyangi et al., 2011). Evaluation of repetitive element methylation and activity may also be pertinent in assessing the viability of pluripotent cells for therapy, as maintained retrotranspositional potential could lead to deleterious insertional mutations over time. Based on these new and emerging roles for SINEs in the genetic and epigenetic regulation of the cell, we decided to compare the methylation profiles of our samples with regards to SINEs specifically. What we saw was the same pattern of methylation for the promoters, the female cell lines showed similar low levels of CpG methylation, resembling the Dnmt knock out cells and the ICM, while once again the males exhibited a more hypermethylated profile (Figure 4)

LINEs (long interspersed nuclear elements) are autonomous counterparts of SINEs. While SINEs are generally located in gene rich areas of the genome; LINEs tend to be located in gene poor regions (Lander et al., 2001). With regards to overall nuclear location, LINEs are found mostly in the transcriptional dormant periphery while SINEs are mostly located in the more active center (Guelen et al., 2007). Why these two elements occupy vastly different genomic compartments of the genome is unclear, especially given the intimacy with which the two are linked. LINEs and SINEs are able to replicate and insert themselves into the genome via retrotranposition. In this process, the retroelements are transcribed by RNA polymerase III, the RNA strand is then converted back to DNA via a reverse transcriptase, and an endonuclease facilitates the insertion of this DNA into the genome (Moran et al., 2006). LINEs contain the DNA sequence which codes for the reverse transcriptase and endonuclease and SINEs do not, making them autonomous. However SINEs do share homology at sequences recognized by essential retrotransposition machinery, and therefore are able to hijack LINE activity to their benefit (Ohshima et al., 1996). LINEs make up ~20% of the mammalian genome, in humans there are over half a million LINE element insertions, though only a small fraction of these are actually active in humans

(Brouha et al., 2003). As stated, LINEs tend to be restricted to gene poor, why this is,

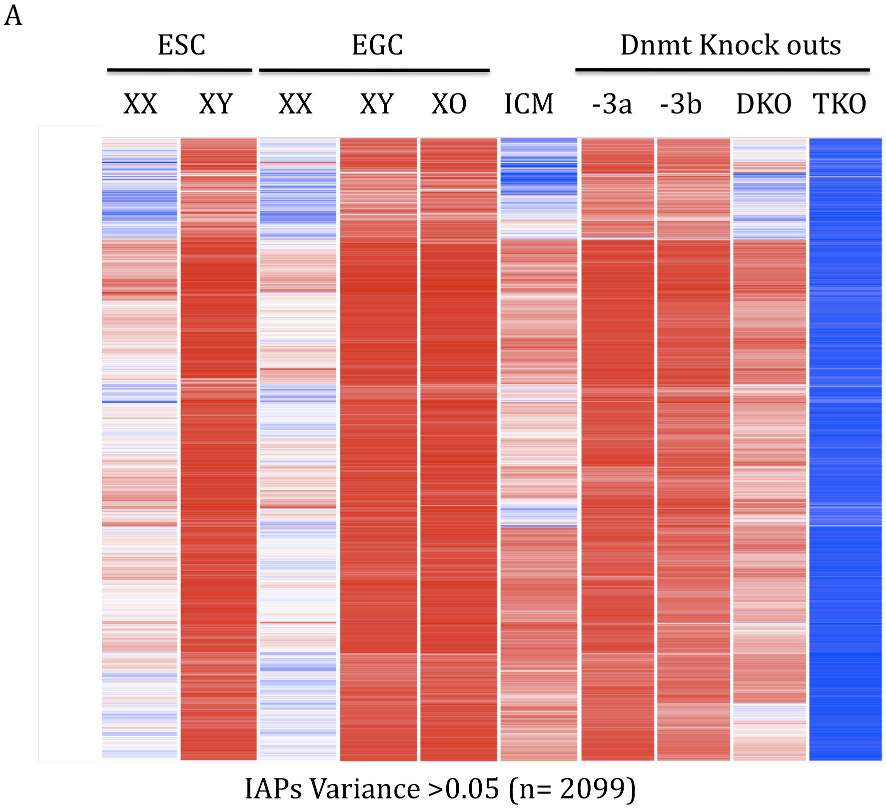


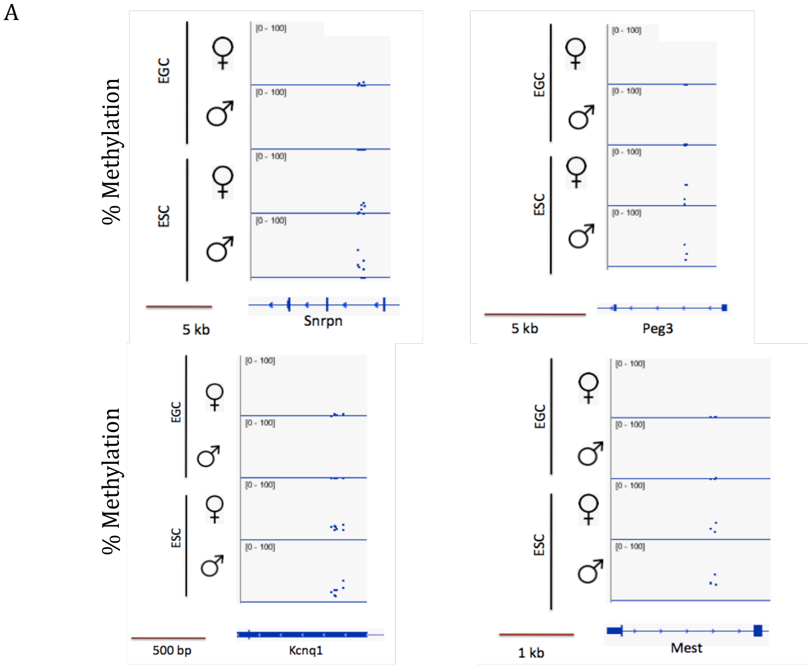
Figure 6. Methylation profile of IAPs in EGCs and ESCs showing variance >0.05 IAPs exhibit hypomethylation in female cells in contrast to males. Dnmt 3a/b double knockout cells and ICM samples also exhibit differential hypomethylation. XO samples resemble male samples. Much like the LINEs, female cells more closely resemble Dnmt 3a/b double knockouts than the ICM.

is unclear, but it may be related to their size. While the insertion of the shorter SINE element can sometimes lead to beneficial mutations to the cell, whether through alternative splicing or enhancer activity, insertion of the longer LINE element may, more often than not, be deleterious. Like SINEs, methylated LINEs are demethylated and reactivated during gametogenesis and early development (Brouha et al., 2002; Luning et al., 2003). Analysis of LINEs in our samples once again showed a similar pattern to the promoter and SINE methylation. (Figure 5)

Intacisternal A particles (IAPs) belong to the endogenous retrovirus family (ERV) class II, and, much like LINEs, are capable of copying and inserting themselves throughout the genome (Maksakova et al., 2006; Kuff et al., 1988). The exact copy number of IAPs in the mouse genome is unclear as they are highly polymorphic as a reflection of their strong activity; as such and they are constitutively repressed to prevent their unchecked expansion (Zhang et al., 2008). Even during the drastic reprogramming that takes place in PGCs, cells carefully maintain strict IAP methylation (Popp et al., 2010). Looking at the methylation profile of our samples it's evident that the male cells are heavily methylated with regards to IAPs, more so than with any other retroelement, and this signature is more or less maintained in the ICM (Figure 6). Here, female lines bare a greater resemblance to Dnmt knockout cells, indicating that their perfuse hypomethylation may not be emblematic of ICM-like regulation, but instead reflects specific and aberrant inhibition of *de novo* methyltransferase function. Unlike SINEs and LINEs where it is hard to distinguish the hypomethylation that we see from the ICM and the Dnmt knock out samples. Our female cells exhibit a global loss of methylation not seen *in vivo* once again raising the question of the viability of these cells for therapeutic applications.

Imprints

The close resemblance of female cells to the Dnmt knock out cells for IAP methylation raises the question of whether females resemble a blank slate emblematic of true pluripotency, without restrictive marks seen even in the ICM, or are simply a result of misregulated methylation. We decided to look at the ICRs of



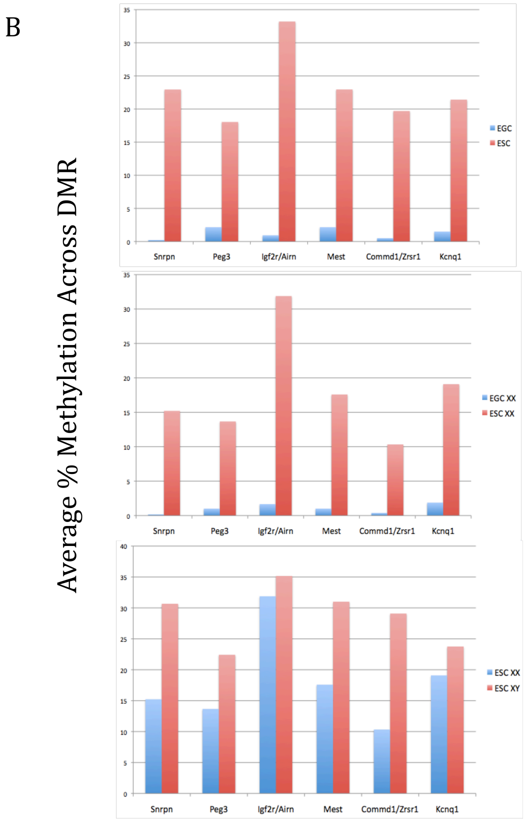
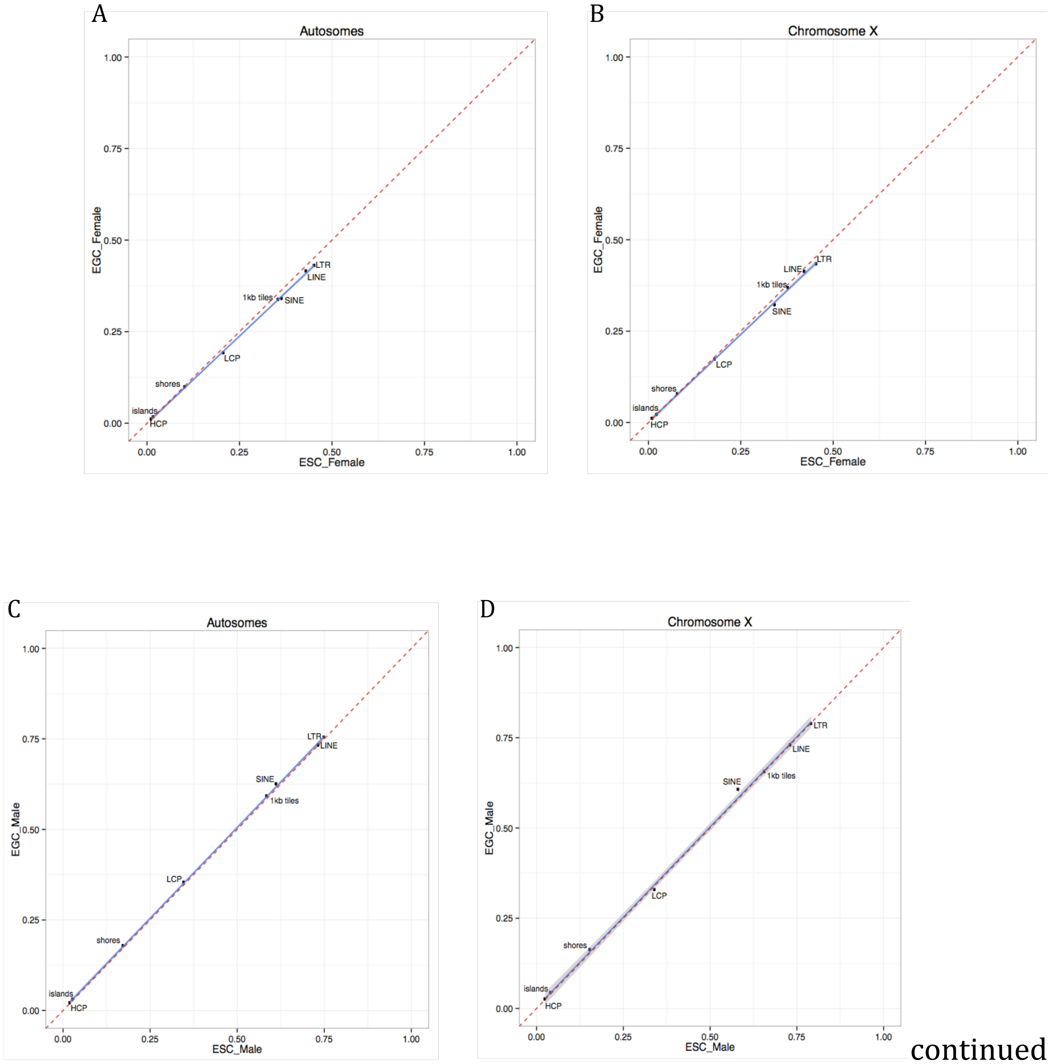


Figure 7. Methylation of DMRs in ESCs vs. EGCs. (A) IGV browser data visualization of the methylation status of select DMRs in ESCs and EGCs. The hypomethylation of female ESCs does not apply to imprinted regions. (B) Quantitative analysis of select DMRs between EGCs and ESCs.

the ESCs and EGCs, both to see if these signatures could distinguish the two cell types (as is expected given that imprints are selectively erased in the germline but not the early embryo) as well as to test if even these signatures were aberrant in female lines. We found that only EGCs are hypomethylated at ICRs, regardless of sex making this the only signature that distinguishes the cells derivation origin vs. their sex. What this may suggest is that imprints are maintained specifically by Dnmt1, while repetitive elements may require persistent Dnmt3a/b activity(Figure 7 A, B).



continued

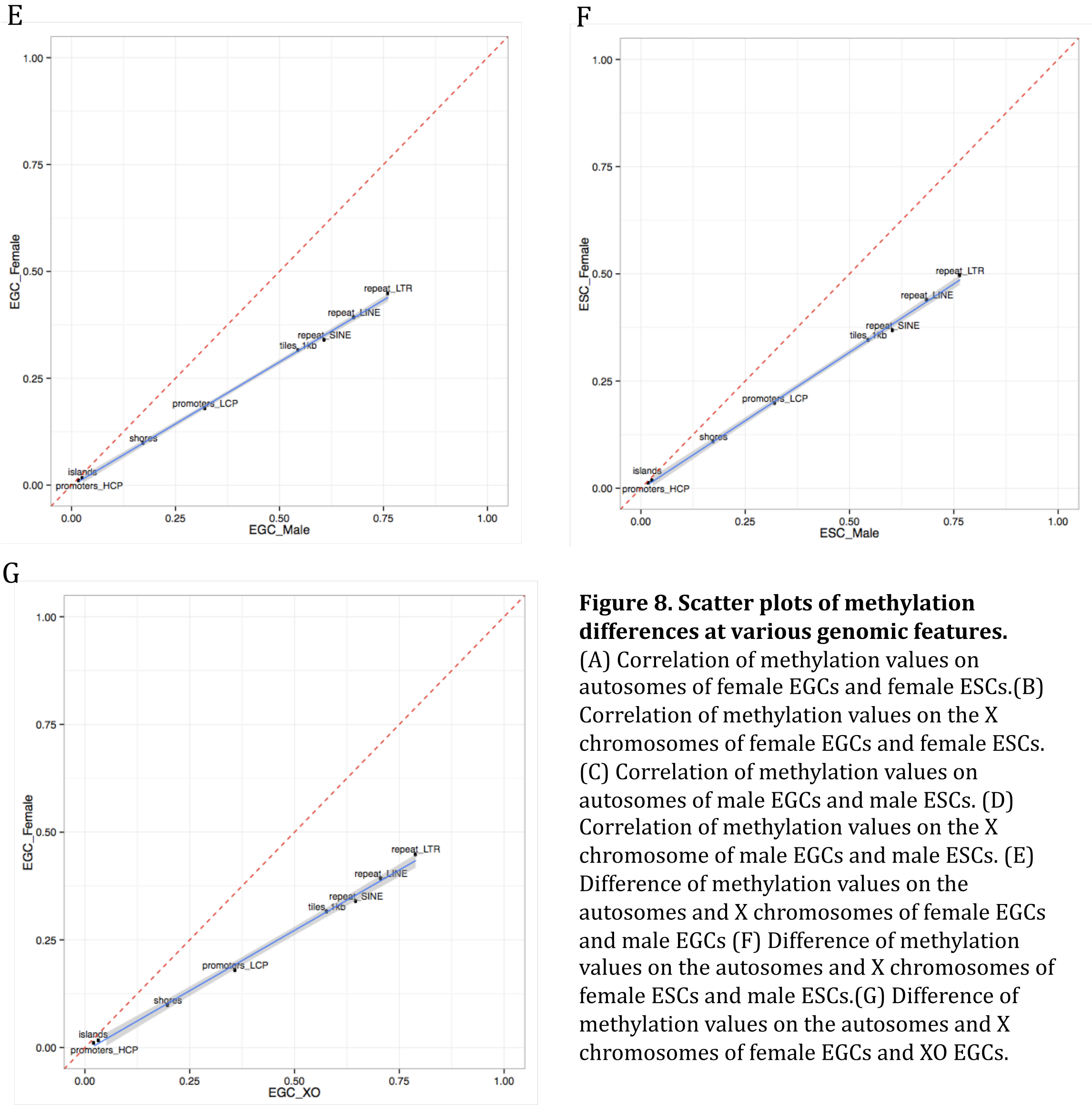
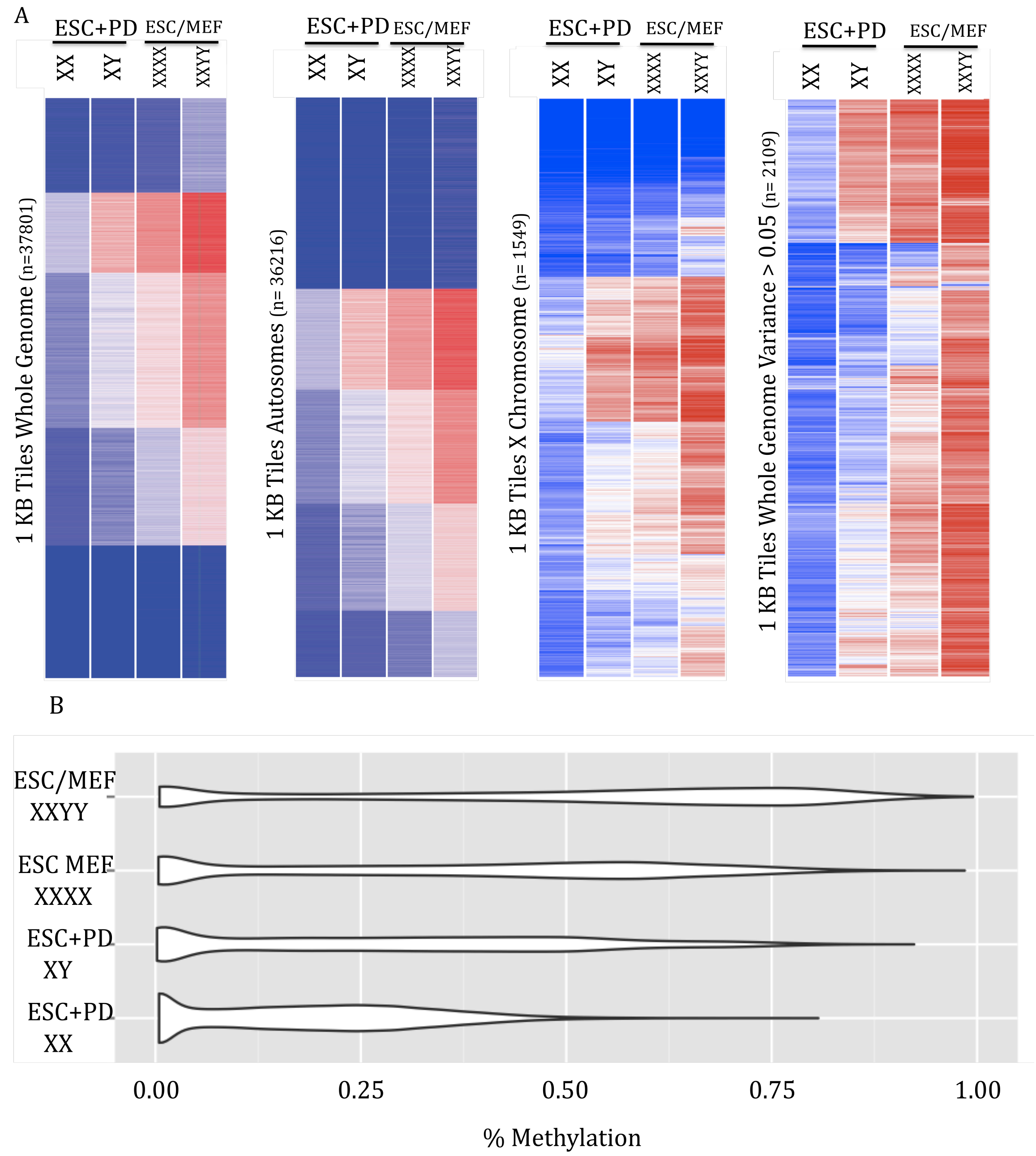


Figure 8. Scatter plots of methylation differences at various genomic features. (A) Correlation of methylation values on autosomes of female EGCs and female ESCs.(B) Correlation of methylation values on the X chromosomes of female EGCs and female ESCs. (C) Correlation of methylation values on autosomes of male EGCs and male ESCs. (D) Correlation of methylation values on the X chromosome of male EGCs and male ESCs. (E) Difference of methylation values on the autosomes and X chromosomes of female EGCs and male EGCs (F) Difference of methylation values on the autosomes and X chromosomes of female ESCs and male ESCs.(G) Difference of methylation values on the autosomes and X chromosomes of female EGCs and XO EGCs.

X chromosome dosage

It is well understood that sex determination in animals is determined by the presence or absence of a Y chromosome. An XXY mammal will develop as a male due to the sex-determining region Y protein (SRY) (Wallis et al., 2008). Alternatively, XO mammals will develop into a female (Uematsu et al., 2002). In insects, such as Drosophila melanogaster, sex is determined in a similar fashion to mammals,


A
ESC+PD
ESC/MEF
XX
XY
XXXX
XXYY
1 KB Tiles Whole Genome (n=37801)
1 KB Tiles Autosomes (n= 36216)
1 KB Tiles X Chromosome (n= 1549)
1 KB Tiles Whole Genome Variance > 0.05 (n= 2109)
B
ESC/MEF XXYY
ESC MEF XXXX
ESC+PD XY
ESC+PD XX
0.00
0.25
0.50
0.75
1.00
% Methylation

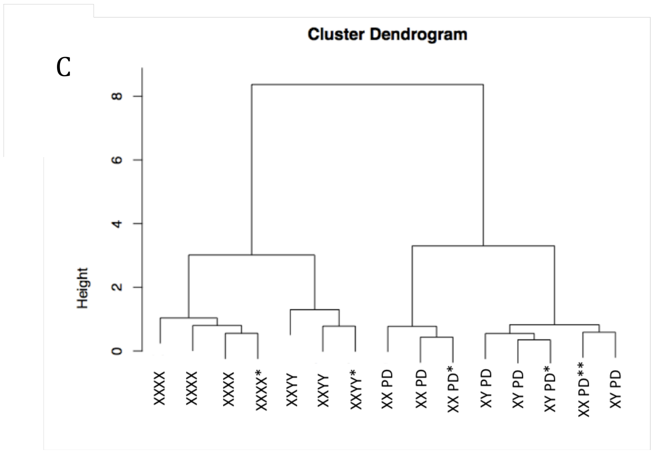
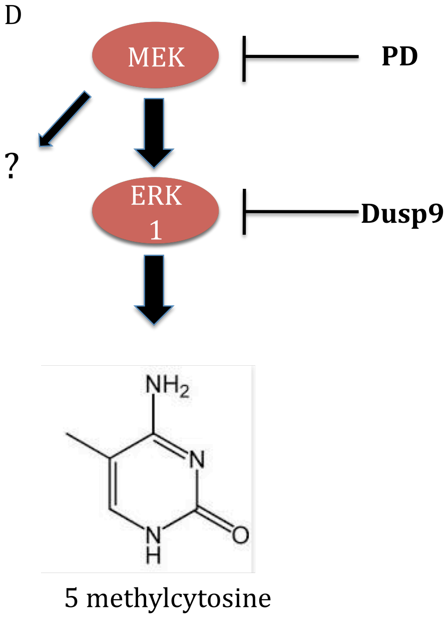
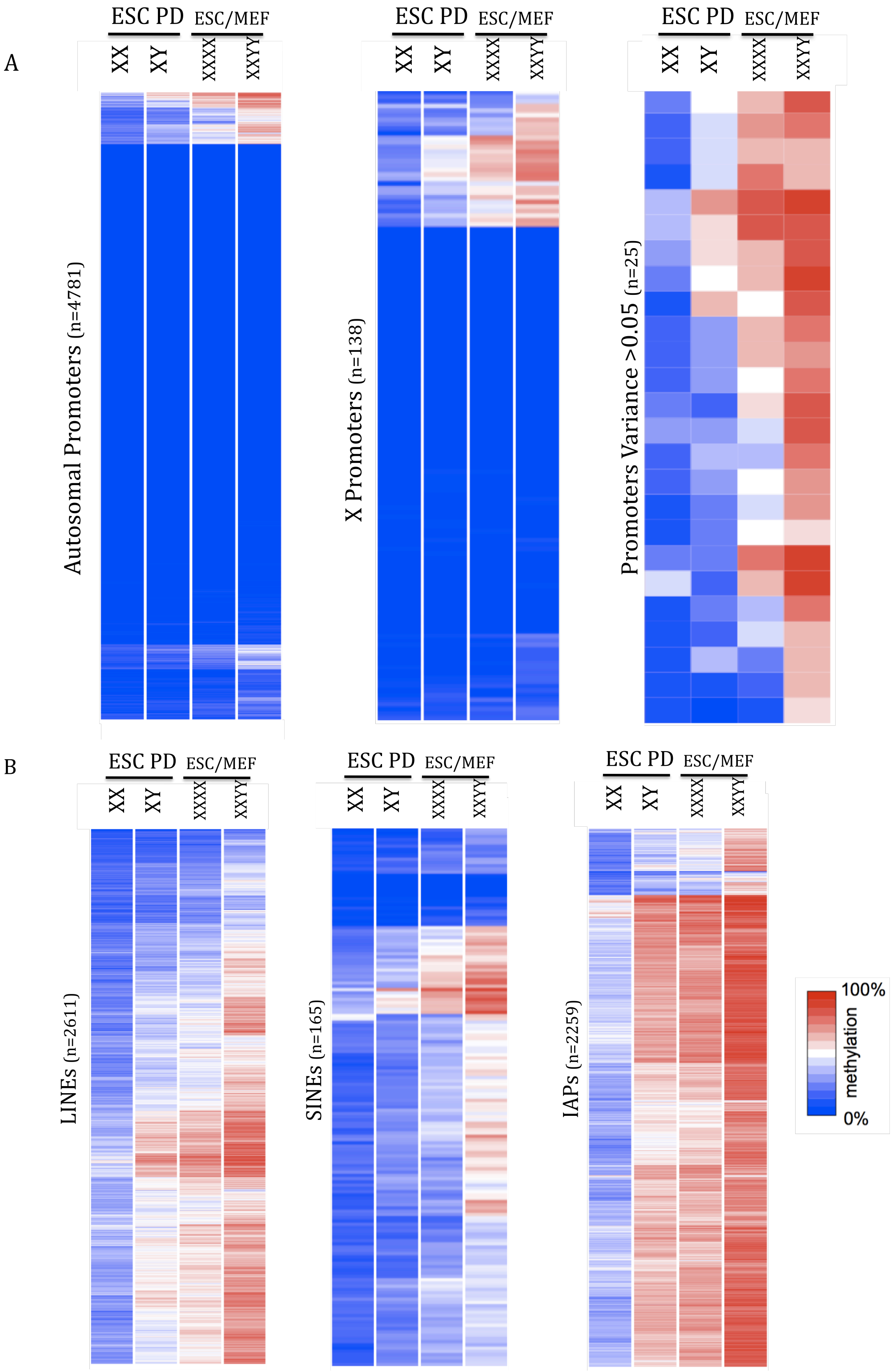


Figure 9. Mek1 Kinase inhibition and X dosage effects. (A) Whole genome methylation profiles of ESCs cultured with the MEK inhibitor PD98059 and ESCs fused with MEFs. (B) Violin plots of total genome methylation. (C) Hierarchical clustering data. * pooled values. **Outlier due to possible mislabeling. (D) Possible model for regulation of DNA methylation via the MEK/ERK pathway.

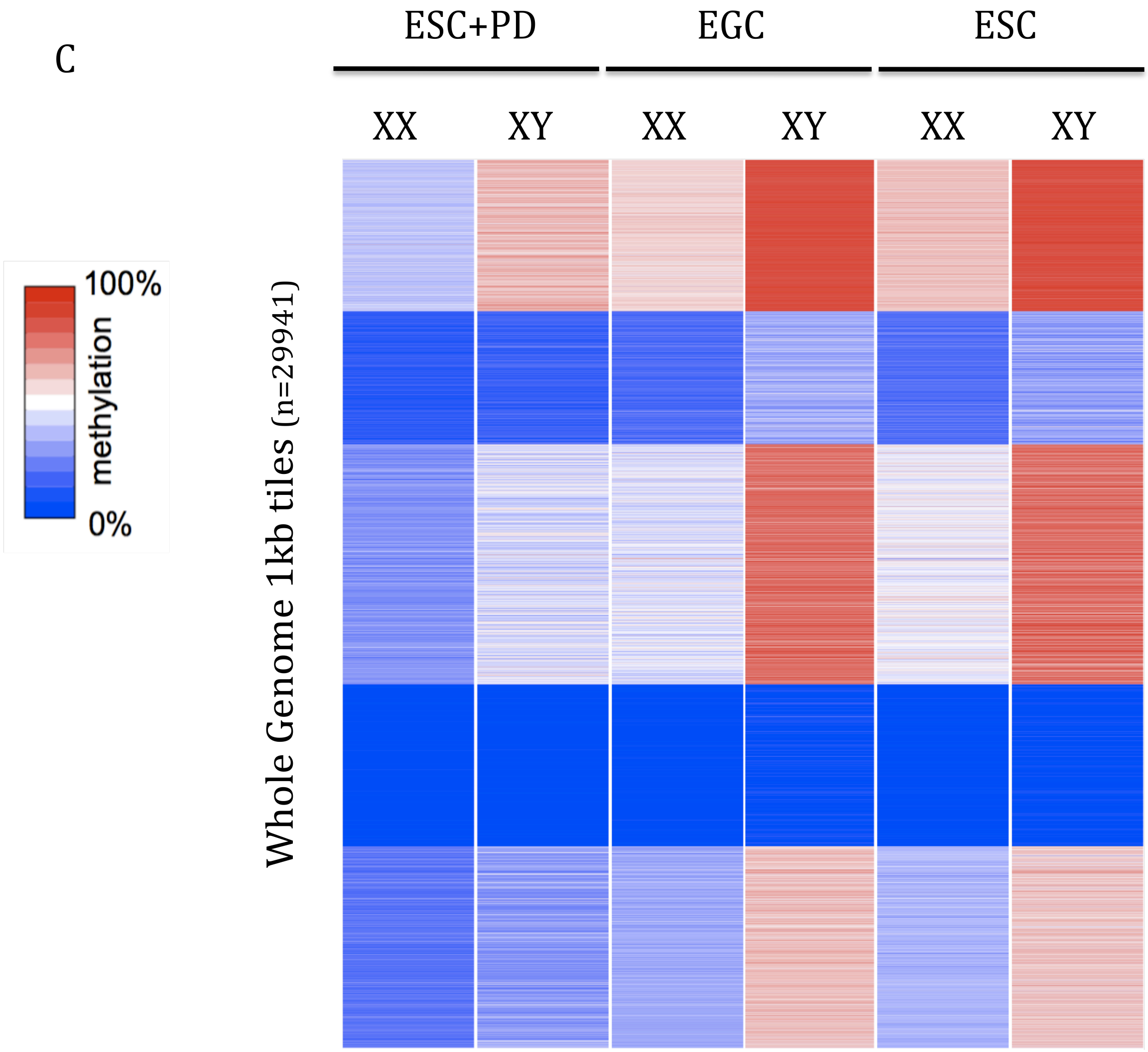


females are XX and males are XY. However the mechanism of sex determination is different. While in mammals the Y encodes the sex determining factor, in flies the determining factor is the ratio of X chromosomes to autosomes. In females, 2 Xs will result in 1:1 X chromosome to autosome ratio, whereas in males XY will result in a 1:2 ratio (Bridges et al., 1921). The fact that our XO samples have a methylation profile similar to our XY samples (Figure 8) suggests that the X chromosome dosage and not the presence of the Y chromosome determines this apparently sex-specific difference in methylation. To investigate this hypothesis more thoroughly, we decided to determine whether the ratio of X chromosomes to autosomes was a determining factor in the hypomethylated phenotype specifically by assessing the methylation profiles of fused cells (Unpublished data, Figure S5b). The data appears to point at an anti-correlation between copy number of X chromosomes and the level of global DNA methylation. XXYY cells are hypermethylated and XXXX cells are hypomethylated, however, XXXY cells exhibit an intermediate level of methylation. This appears to be indicative of a gene dosage affect (Veitia et al., 2013).

One factor identified by our collaborators as a possible regulator of global methylation in ESCs was Dusp9 (unpublished data Figure S5a). Dusp9 is a gene located on the X chromosome and is a known inhibitor of ERK signaling (Urness et al., 2008). ERK signaling has been shown promote differentiation in ESCs (Kunath et.al., 2007), and inhibition of the MEK1 kinase, an upstream regulator of ERK signaling, allows for greater efficiency in the derivation of ESCs *in vitro* (Buehr et al., 2003). Over expression of Dusp9 results in reduced global methylation in male ESCs (unpublished data Figure S5a). With this in mind we decided to compare the methylation profiles of cells cultured with the MEK inhibitor PD98059 (PD) with ESCs fused with MEFs resulting in cells with either four or two X chromosomes (Figures 9 and 10). Female ESCs cultured with PD exhibit an increased level of hypomethylation across all genomic features, compared with male cells cultured with PD. ESC/MEF XXXX samples exhibit a level of methylation similar to that of male ESCs cultured in PD, however, XY PD cells are still slightly less methylated (Figure 9b). XXYY samples show the highest levels of methylation by comparison. These results indicate a larger role for Mek1 in the maintenance of DNA methylation, and suggests hitherto unidentified targets downstream of MEK1 in addition to ERK.


A
ESC PD
ESC/MEF
XX
XY
XXXX
XXYY
Autosomal Promoters (n=4781)
X Promoters (n=138)
Promoters Variance >0.05 (n=25)
B
LINEs (n=2611)
SINEs (n=165)
IAPs (n=2259)
100%
methylation
0%

C


ESC+PD
EGC
ESC
XX
XY
XX
XY
XX
XY
100%
methylation
0%
Whole Genome 1kb tiles (n=29941)

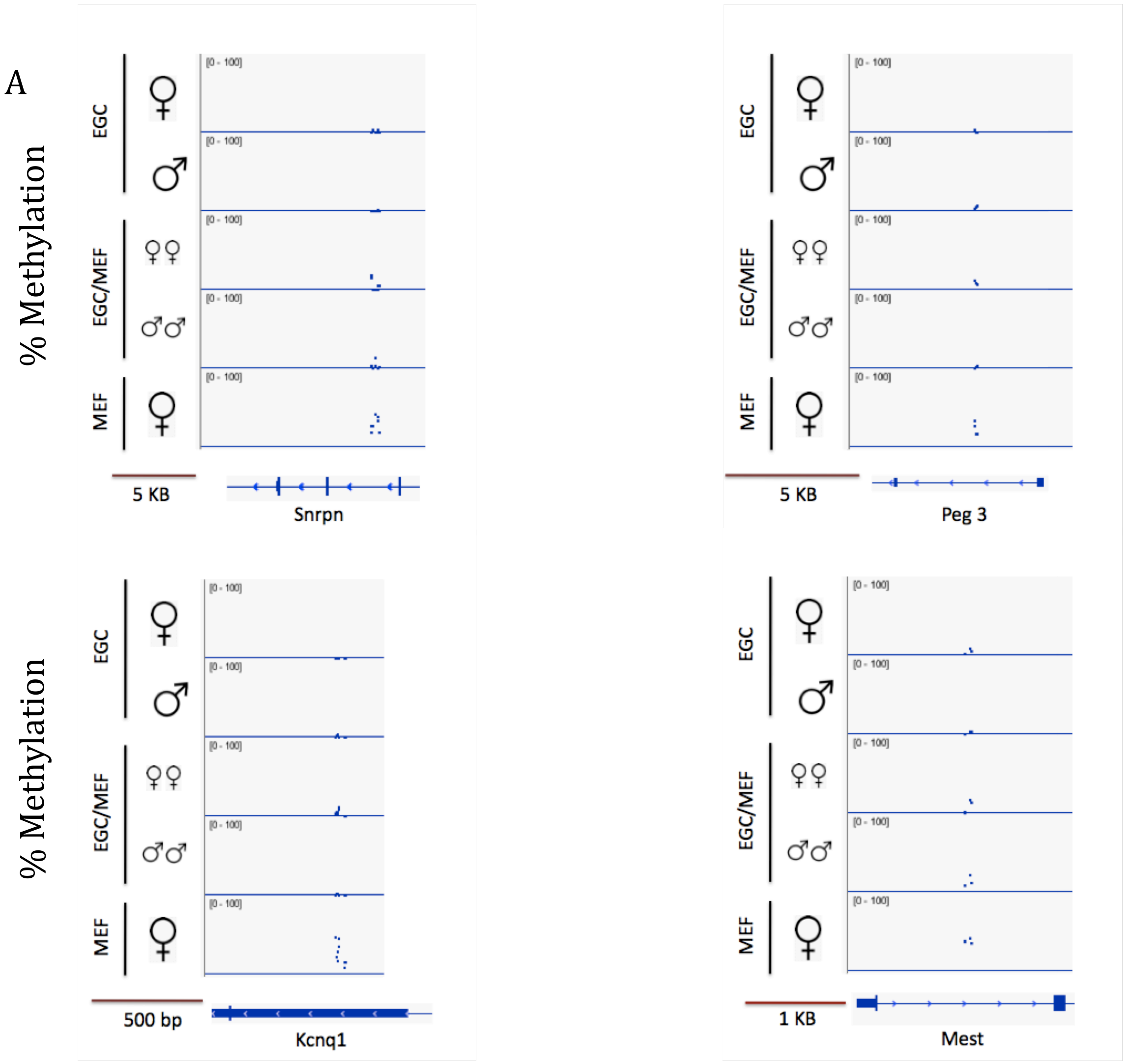
Continued.

D

Cell Type	Sex	n
ES PD	XX	3
“	XY	3
ESC/MEF	XXXX	3
“	XXYY	2

Figure 10. DNA methylation levels across various genomic features in both fused and PD cultured cells. (A) Methylation of autosomal promoters, X promoters, and promoters showing >0.05 variance. (B) Methylation of LINEs, SINEs, and IAPs. (C) Comparison of PD cultured cells with ESCs and EGCs (D)) Number of each cell line replicate pooled together.

To ensure that our EGC samples were fully functioning, they were fused with MEFs to determine their ability to demethylate the somatic genome with specific attention to ICRs. (Figure. 11, S7) Comparison of fused EGC/MEFs with unfused MEFs showed a marked reduction in DNA methylation at ICRs. Additionally it appears that, while female EGCs and ESCs, exhibit similar methylation profiles, EGC fusion cells are slightly less methylated than ESC fusion cells (Figure 12). Most importantly, however, are the results from the XXYY fusion samples. Male EGCs/ESCs fused with male MEFs exhibit a more hypomethylated genome than male EGCs, ESCs and MEFs alone (Figure.12) This hypomethylation may be attributed to the prescence of an additional X chromosome, strengthening the case for the role of Dusp9 in maintaining a hypomethylated genome. Bizarrely, Female EGCs/ESCs fused with female MEFs are more methylated than female EGCs/ESCs alone. They are still hypomethylated when compared to the males (Figure. 12).


A
% Methylation
% Methylation
EGC
EGC/MEF
MEF
♀
♂
♀♀
♂♂
[0 - 100]
5 KB
Snrpn
5 KB
Peg 3
500 bp
Kcnq1
1 KB
Mest

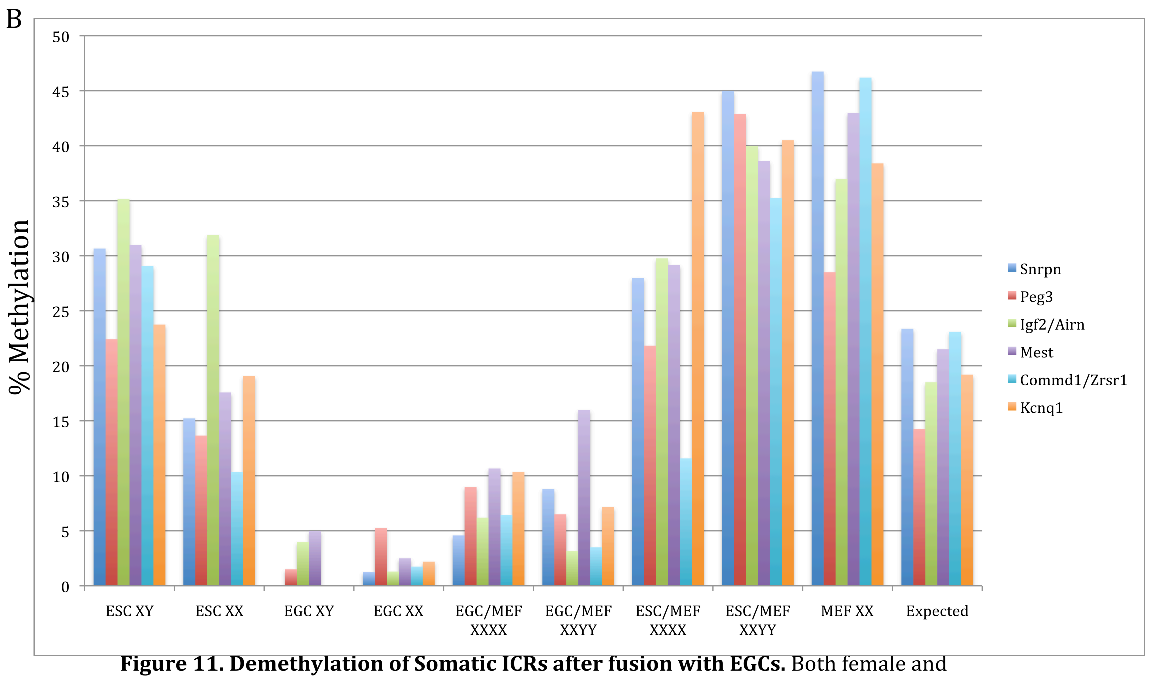
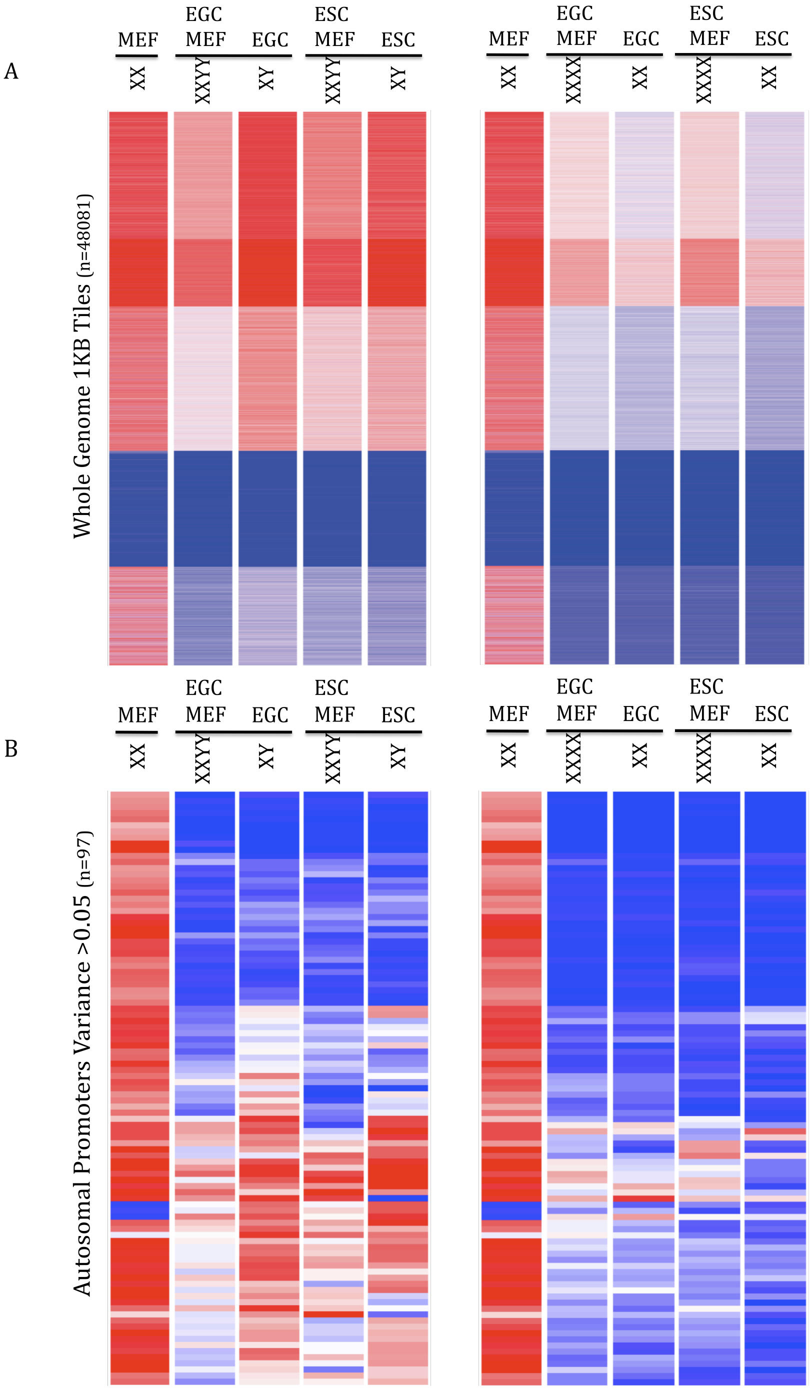
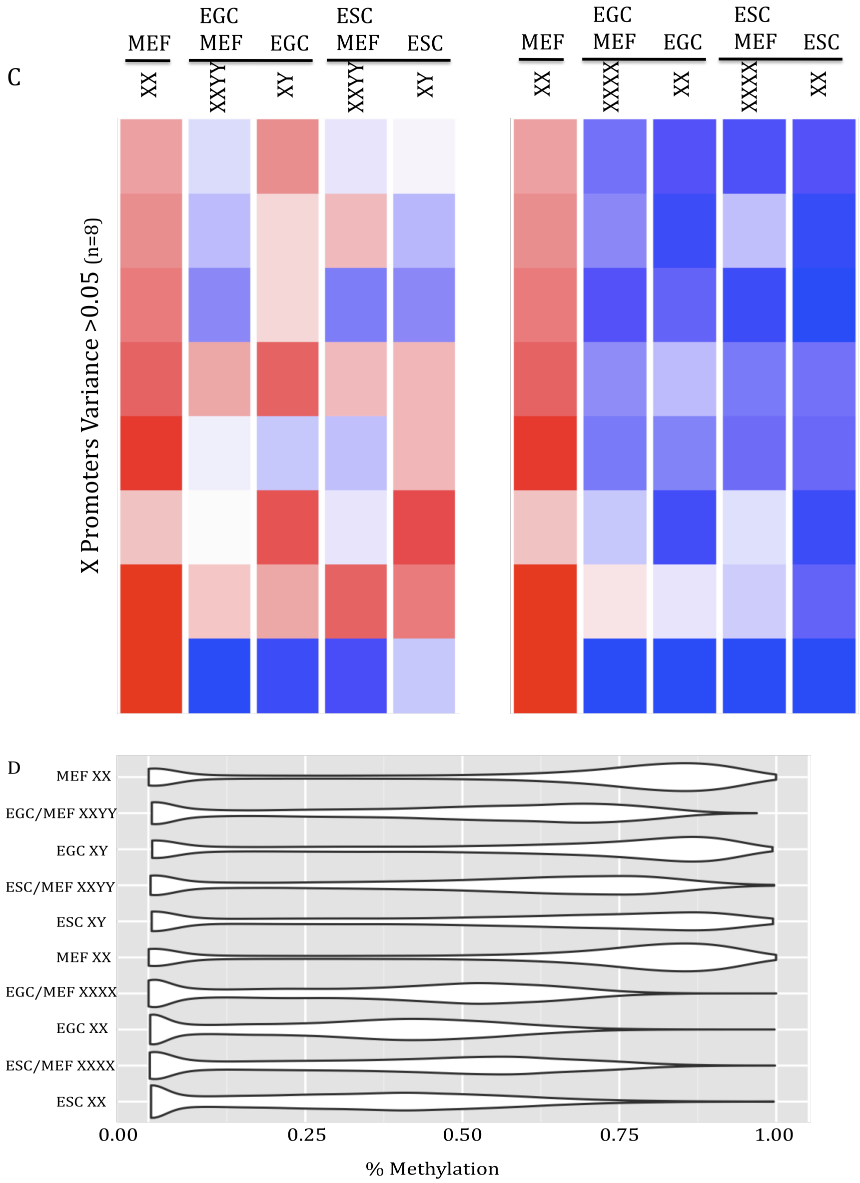


Figure 11. Demethylation of Somatic ICRs after fusion with EGCs. Both female and male EGCs exhibit loss of DNA methylation at ICRs. Methylation is drastically reduced at the ICRs in MEFs fused with EGCs (A) IGV browser data visualization of the methylation status of select DMRs in male and female EGCs, fused EGC/MEF samples, and MEFs. (B) Quantitative analysis of the the methylation status of DMRs in both fused and unfused cells. "Expected" columns indicate where EGC fusion levels would be in the absence of demethylation activity.

Taken together all this data points to a striking and bizarre methylome difference between male and female pluripotent cells that is globally apparent for a number of different genomic features (Figure 8, Figure S6). Comparing male ESCs with male EGCs we can see an almost perfect alignment for all genomic features (Figure 8). The female ESCs and EGCs are not as perfect a match, with the EGCs showing some hypomethylation in the repeat elements compared with ESCs, but this difference is minimal and may reflect variation between lines (Figure 8). The most informative data comes from the comparison of female to male pluripotent cells. As expected, there is little difference in high CpG density promoters (HCP). HCPs are usually associated with housekeeping genes and genes involved developmental regulation (Saxonov et al., 2006). Regardless of sex, all pluripotent cells would be expected to share this feature. Differences between sexes become apparent when promoters with Low CpG density (LCPs) are investigated. LCPs are more commonly found at tissue specific genes, and previous work has shown that the majority of LCPs are methylated in a given cell type. Why then we should see differences between the male and female lines is unclear. The same subtle differences can be seen in the “shores” of CpG islands, which are generally more dynamic and associated with transcriptional changes (Figure 8)(Irizarry et al., 2009). The largest differences are seen in the retroelements of LINE, and SINE and long terminal repeats (LTRs) classes. Why differences between pluripotent cells of either sex would be predominantly found for these features is unclear, and their proximity to the global description from 1 KB tile data appears to be indicative of DNA methylation misregulation. Previous work has shown that repetitive elements exhibit an increased sensitivity to demethylation in the absence of Dnmt3 a/b (Arand et al., 2012) .


A
MEF
EGC
MEF
EGC
ESC
MEF
ESC
XX
XXYY
XY
XXYY
XY
MEF
EGC
MEF
EGC
ESC
MEF
ESC
XX
XXXX
XX
XXXX
XX
Whole Genome 1KB Tiles (n=48081)
B
MEF
EGC
MEF
EGC
ESC
MEF
ESC
XX
XXYY
XY
XXYY
XY
MEF
EGC
MEF
EGC
ESC
MEF
ESC
XX
XXXX
XX
XXXX
XX
Autosomal Promoters Variance >0.05 (n=97)



E

Cell type	Sex	n
MEF	XX	1
EGC	XX	1
EGC	XY	1
ESC	XX	1
ESC	XY	1
EGC/MEF	XXXX	3
EGC/MEF	XXYY	2
ESC/MEF	XXXX	3
ESC/MEF	XXYY	2

Figure 12. EGCs/ESCs fused with MEFS exhibit similar methylation profiles. (A) Genome Scale 1KB tiling of MEFs, EGCs, ESCs and fusion samples. (B) Autosomal promoters showing >0.05 variance. (C) X chromosome promoters showing >0.05 variance. (D) Violin plots of “global” DNA methylation. (E) Table of cells, their sex, and amount pooled. Single EGC and ESC samples correspond to the cell lines used in the fusion experiments.

Discussion

Based on these results it is clear that the methylome of female cells is drastically different than that of the males. What does this mean? Are female stems cells more pluripotent or closer to a "naïve" ground state of pluripotency? Does this hypomethylation confer any benefit to the cell, or is it hazardous to development. Recent work by Head et al. shows that in female ESC, the presence of two active X chromosomes results in a range of *in vitro* differences, including delayed differentiation of ESCs in culture, which they attribute to a global hypomethylation similar to what we see in our samples (Edda et al., 2014). They report that inhibition of the MAPK and Gsk3 pathways (the same pathways inhibited by 2i media), as well as the activation of the Akt pathway, are responsible for this hypomethylation phenomenon, which is supported by the observation that culturing pluripotent cells in 2i media results in a hypomethylated genome (Wu et al., 2014). Additionally, they report a specific repression of Dnmt3a/b in female cells downstream of PRDM14, which is usually inhibited by the Gsk3/erk pathways. They also report that female cells cannot differentiate until after one X chromosome has been inactivated, and that this differentiation barrier can be lifted via ectopically induced X chromosome inhibition. They conclude that X inactivation and differentiation are functionally linked, and that this relationship acts as a safety mechanism whereby development cannot progress until the exit from pluripotency is assured by random X inactivation. The need for this regulation can be seen in the case of Beckwith-Wiedeman syndrome, a disease that results from the improper methylation of the imprinted KvDMR1 locus. This disease affects monozygotic twins at a much higher rate than the rest of the population, and of these twins it is seen nearly entirely in females (Lubinsky et al., 1991). It is speculated that, as with normal development, all X chromosomes are active, however in monozygotic twins there may be an extended period when all Xs' are active during cell development resulting in these epigenetic abnormalities. Persistent expression of two X chromosomes in mice results in embryonic death at approximately 10 dpc (Takagi et al., 1990).

From its inception, the goal of this project was to identify the factors present in EGCs that enabled them to reprogram somatic cells in such a fashion that ICR were also targeted for demethylation in addition to their shared ability with ESCs to overhaul the methylome of the somatic cell. In order to do this, we collaborated with the Lab of Konrad Hochedlinger to better understand differences between EGCs and ESCs that may provide informative EGC-specific reprogramming candidates. While we focused on the methylation profiling of the samples, their group investigated the transcriptional dynamics of the two cell types. Similar to our results, RNA-sequencing data of EGC and ESC samples clustered in a manner similar to what we see with DNA methylation, namely that they cluster based on sex and not by cell type (unpublished data, Figure S3). While novel, these observations do not provide clear insight into the ability of EGCs to demethylate ICRs, though it does reinforce the important role X inactivation plays in development.

From screening for transcriptional differences between cells of either sex, one factor that did stand out was Dusp9, which is specifically reduced in male cells compared to female cells (unpublished data, Figure S4). Over expression of a doxycycline inducible Dusp9 resulted in a significantly reduced level of methylation in male cells, however the level of hypomethylation was not the same as in wild type female cells, indicating a role for other factors on the X chromosome (unpublished data, Figure S5).

While no defining factor has yet been recovered with regards to the different reprogramming capabilities of EGCs and ESCs, the results of this study have put to rest some of the exaggerated reports previously made about EGCs and highlight the negative impact investigation on only a small cohort of sample lines can have on data interpretation. From the first Tada paper to the more recent research by Amanda Fisher, EGCs have always been regarded as a unique pluripotent cell type with a genome miraculously devoid of DNA methylation (Piccolo et al., 2013; Tada et al., 1997). Our results suggest that these groups may have been inadvertently using female EGC lines, or that sex differences between ESC and EGC samples may account for the majority of their observations.

There are two possibilities that may explain differences in ICR methylation between EGCs and ESCs. Either a factor is present in EGCs which specifically targets and demethylates these loci, or there is a factor absent from EGCs which is responsible for the protection of these regions. Zfp57 at a hexanucleotide consensus sequence TGCCGC present in ICRs, resulting in corecruitment of its partner TRIM28 and protection of the bound region from demethylation. This is only possible when the third cytosine is modified with a methyl group; if cytosine is unmodified or hydroxymethylated, Zfp57 bind affinity is greatly reduced (Liu et al., 2012).

Additionally, Zfp57 knockout embryos lose methylation ICRS (Li et al., 2008). However, side-by-side global comparison for all the different Zfp proteins showed that EGCs and ESCs were similar in their expression levels, indicating that the regulation of ICRs is unlikely to operate in the transcriptional regulation of methylation-dependent binding proteins (Corsinotti et al., 2013).

Perhaps genome wide profiling is not the only approach one can take to identify the differences between these cell types. Utilizing an alternative biochemical approach, I was able to identify differences between ESCs and EGCs by assessing the differential DNA binding properties of EGC and ESC proteins in cell lysates. Using DNA strands that contained either methylated, hydroxymethylated, or unmodified cytosines, differential binding factors are evident between EGC and ESCs as well as for cytosine modifications (Figure 13. This approach, coupled with Mass Spectrometry to identify differential binding partners, may be a useful next step in identifying pertinent factors related to ICR regulation in EGCs.

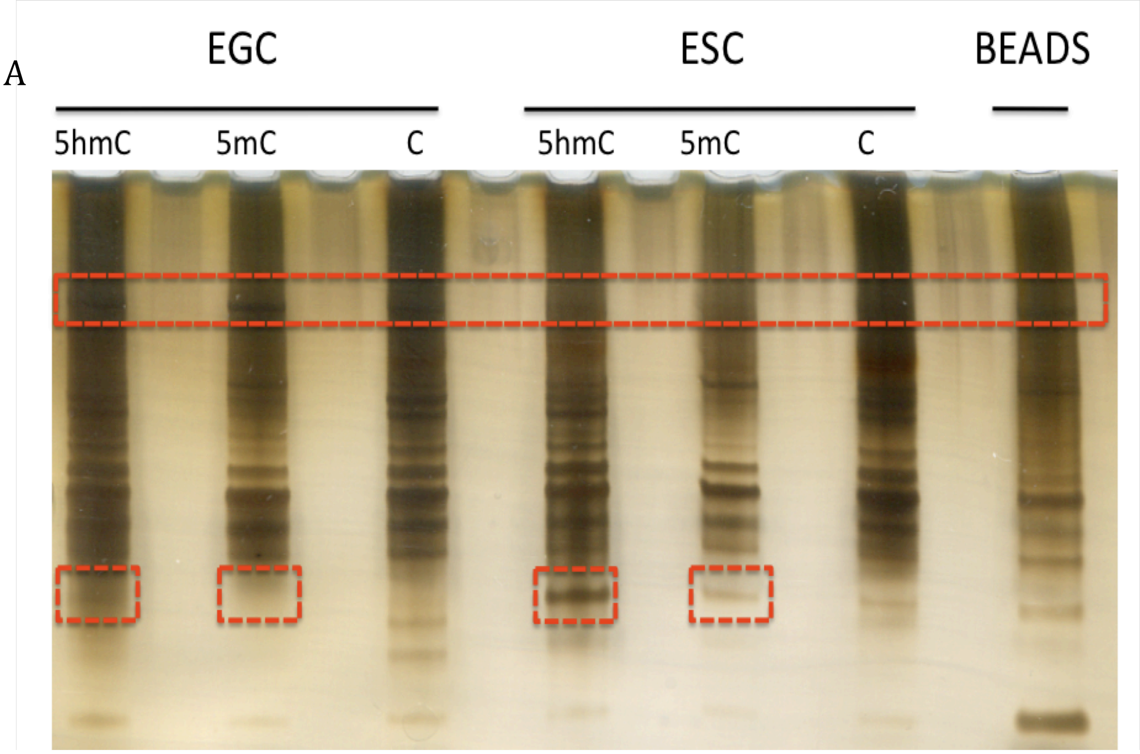


Figure 13. Differential binding of EGC and ESC proteins to modified DNA bait. Silver stain of proteins eluted from modified DNA oligos. EGCs exhibit bands missing from ESCs (red rectangle). Differential binding can be observed between 5hmC and 5mC (red boxes).

METHODS

Cell Isolation

Embryonic stem cells were isolated as described in Meissner et al., 2009 C57BL/6 females were mated with 129 males. At 3.5dpc females were sacrificed and blastocysts were isolated from the uterine horn. After collection of blastocyts the zona pellucida is dissolved by incubating in acid tyrodes. Blastocysts with the zona removed a placed into plates with a MEF feeder layer in ESC media (ES cell medium: 425 ml DMEM (Gibco/Invitrogen, knockout #10829-018), 75 ml fetal bovine serum (FBS, heat inactivated, 56°C for 30 min, hyclone characterized #SH30071.03), 5 ml non-essential amino acids, 5 ml penicillin/streptomycin, 5 ml glutamine, 4 ml betamercaptoethanol, 50 ml LIF ESGRO) conditioned with PD98059. After 4-6 days an ICM outgrowth can be seen. This outgrowth was trypsinized and transferred to new MEF coated plates. Within 3 days colonies are visible.

Embronic germ cells were isolated as described in Durcova-Hills et al., 2008 C57BL/6 females were mated with 129 males. At 11.5dpc females were sacrificed and embryos are removed and placed in ice cold PBS. The genital ridges of the embryos are removed and trypsinized. Each individual genital ridge is divided across two wells of a standard four well plate with EGC media (425 ml DMEM (Gibco/Invitrogen, knockout #10829-018), 75 ml fetal bovine serum (FBS, heat inactivated, 56°C for 30 min, hyclone characterized #SH30071.03), 5 ml non-essential amino acids, 5 ml penicillin/streptomycin, 5 ml glutamine, 4 ml betamercaptoethanol, 50 ml LIF ESGRO, 25 ng/ml FGF-2). After 6-8 days colonies are visible.

Reduced Representation bisulfite Sequencing (RRBS)

The application of RRBS allows us to create genome "scale" methylome maps of cell. The term genome scale refers to the ability to select only pertinent portions of the genome for sequencing thereby reducing costs. Libraries were prepared as described in Boyle et al., 2012[i] DNA was isolated from EGCs and ESCs taking care to remove proteins and excess RNA which can have adverse effects on bisulfite conversion. The DNA is then digested with the restriction enzyme MspI. MspI recognizes the sequence CCGG and cuts between the two cytosines. Digesting with this enzyme ensures that each fragment will contain at least two informative CpGs, one at each end. After filling in the 3' end of the fragment, A tailing is performed to allow for adapter ligation. Methylated adapters are utilized as they will not be converted in the bisulfite conversion step. Bisulfite conversion will convert unmethylated cytosines to uracil, while leaving methylated cytosines untouched. After illumina sequencing fragments are realigned to a reference genome thereby informing us of the methylation status of a genomic region.

Region methylation averages were calculated using a weighted average of CpGs that were well-covered in at least 75% of selected samples. Regions in which fewer than 5 CpGs were not used in downstream analysis. The genome was tiled into non-overlapping 1Kb tiles for global analysis. Promoter regions were defined as the 2Kb region centered at the transcription start site of RefGene genes. Hierarchical clustering was performed on region methylation values using euclidean distance and Ward's method for clustering. Variable regions included regions with a variance of > 0.05 across the methylation values for all samples. Plots were created using custom R scripts.

DNA pull down assay

1 Kb biotinylated oligonucleotides of the H19 locus where created using the following primers:

FORWARD-[Biotin~5]GCGATGTACGAGACTTCACT

REVERSE- AATAGAGATTCTATTTTCAT

Three different oligonucleotides were created using dCTPS, d5mCTPs (5-methylcytosines), and d5hmCTPs (5-hydroxymethylated cytosines).

PCR protocol

5μl primers (2.5μl of each forward and reverse) = 0.5μM

5μl 10X Taq buffer

1μl genomic DNA (1ng/μl)

0.1μl of each dNTP =0.4μM

1μl Taq polymerase

38.6μl H_2O

(1μl DMSO improves efficiency)

Protein Isolation

EGCs and ESCs Lysate prep (I use 1.5 mg total): 2x Lysis buffer (150 mM KCl, 25 mM TRIS- HCl pH 7.4, 5mM EDTA, 0.5% NP-40) + 1X protease inhibitor and 0.5mM DTT. Then I add 5 mM MgCl2, 1 ul of RNaseOut. These values are per pull down.

DNA immobilization on streptavidin/agarose beads

50μl of beads were washed with twice with a 2x wash buffer (10 mM Tris, 1mM EDTA, 2.0M NaCl, pH adjusted to 7.0 by addition of HCl).

Biotinylated DNA was added to beads with a 1x wash buffer and allowed to incubate at room temperature for ~15 minutes. Beads were then washed twice with 1X wash buffer

Protein and DNA incubation

Beads are washed three times with an incubation buffer (50mM Tris, 1mm EDTA. 100mM KCl, pH adjusted to 7.0 with HCl. Then add 5% glycerol and 0.1% TritonX100 and 100mM DTT). Cell lysate is added in a 1:1 ratio with the incubation buffer. Lysate and DNA are incubated for 30 minutes at 70 ^{o}C. Taking care to shake every 5 minutes. After incubation the supernatant is removed and the beads are washed 3 times with wash buffer to remove non-DNA-bound proteins.

Gel analysis

5X Laemmli buffer (4ml 1.5M Tris-Cl pH 6.8. 10ml glycerol, 5 ml betamercaptoethanol, 2g SDS (sodium dodecyl sulfate), 1ml 1% bromophenol blue). Add to 1X concentration to washed beads and heat to 50^0C to elute bound proteins. Run samples on premade SDS gel and silver stained to allow visulaiztion of the protein bands.

Bibliography

Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, Jenuwein T, Xu G, Leonhardt H, Wolf V, et al. (2012) *In Vivo* control of CpG and Non-CpG DNA methylation by DNA methyltransferases. PLoS Genet.8:e1002750.

Babaie, Y., Herwig, R., Greber, B. et al (2007). Analysis of OCT4 dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. Stem Cells *25*, 500–510.

Ball, M. P. et al (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nature Biotechnol. 27, 361–368.

Bartke et al (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell *143.3*, 470-84.

Bartolomei, M. S., Zemel, S., and Tilghman, S. M. (1991). Parental imprinting of the mouse *H19* gene. Nature *351*, 153-155.

Bejerano, G., Lowe, C. B., Ahituv, N., King, B., Siepel, A., Salama, S. R., Rubin, E. M., Kent, W. J., and Haussler, D. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature 441,87–90

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209–213.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Blelloch et al (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells *24.9*, 2007-13.

Boyle et al. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome biology (2012) vol. 13 (10) pp. R92

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536–2539.

Buehr, M. and A. Smith, (2003) Genesis of embryonic stem cells. Philos Trans R Soc Lond B Biol Sci, 358(1436): p. 1397–402; discussion 1402.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev. Biol. *210*, 30–43.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255–256.

Bridges C B (1921). Triploid intersexes in *Drosophila melanogaster*. Science. 54:252–254.

Briggs, R. King, T.J. (1953). J. Expol. Zool. *122*, 485.

Brinkman, A. B. et al (2012). Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res. *22*, 1128–1138.

Brons et al (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-5.

Brouha B, Meischl C, Ostertag E, de Boer M, Zhang Y, Neijens H,Roos D, Kazazian HH Jr: (2002) Evidence consistent with human L1 retrotransposition in maternal meiosis I. Am J Hum Genet, 71 :327-336

Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr: (2003)Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci USA,100:5280-5285.

Brun, R. B. and Du Pasquier, L. (Dec. 1975). Lymphocytes of the toad Xenopus laevis have the gene set for promoting tadpole development. Science *26,* 1310-1312.

Campbell, K. H. S., Ritchie, W. A., and Wilmut, I. (1993). Nuclear- cytoplasmic interactions during the first cell cycle of nuclear transfer reconstructed bovine embryos: implications for deoxyribonucleic acid replication and development. Biol. Reprod. *49*, 933–942.

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell *20*, 845–854.

Cattanach, B. M. (1986). Parental origin effects in mice*.* J. Embryol. Exp. Morphol*. 97*, 137–150.

Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells *26*, 1496–1505.

Chambers, I., Silva, J., Colby, D. et al (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chazaud, C., Yamanaka, Y., Pawson, T. and Rossant, J.(2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev. Cell *10*, 615-624.

Clouaire, T. *et al* (2012)*.* CFP1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev*. 26*, 1714–1728.

Coan, P., Burton, G., and Ferguson-Smith, A. C. (2005). Imprinted genes in the placenta. Placenta *26*, S10–S20.

C.V. Dang (2012). MYC on the path to cancer. Cell, *149*, 22–35.

Corsinotti A, Kapopoulou A, Gubelmann C, Imbeault M, Santoni de Sio FR, et al. (2013) Global and Stage Specific Patterns of Krüppel-Associated-Box Zinc Finger Protein Gene Expression in Murine Early Embryonic Cells. PLoS ONE 8(2)

Davis, R. L., Weintraub, H., and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987–1000.

de Sousa Lopes, Chuva et al (2008). X chromosome activity in mouse XX primordial germ cells. PLoS Genet *4.2*, e30.

de Sousa Lopes, SM et al (2004). BMP signaling mediated by ALK2 in the visceral endoderm is necessary for the generation of primordial germ cells in the mouse embryo. Genes Dev. Aug 1;18(15):1838-49.

Durcova-Hills, G., Adams, I.R., Barton, S.C., Surani, M.A., McLaren, A. (2006). The role of exogenous fibroblast growth factor-2 on the reprogramming of primordial germ cells into pluripotent stem cells. Stem Cells *24*, 1441–1449.

Durcova-Hills et al (2008). Reprogramming primordial germ cells into pluripotent stem cells. PLoS ONE *3.10*, e3531.

Dvash, T., Lavon, N., Fan, G. (2010). Variations of X Chromosome Inactivation Occur in Early Passages of Female Human Embryonic Stem Cells. PLoS ONE 5(6): e11330.

Edda G. Schulz, Johannes Meisig, Tomonori Nakamura, Ikuhiro Okamoto, Anja Sieber, Christel Picard, Maud Borensztein, Mitinori Saitou, Nils Blüthgen, Edith

Heard. (2014) The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. Cell Stem Cell - (Vol. 14, Issue 2, pp. 203-216)

Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W. III, Yanagimachi, R., Jaenisch, R (2000). X-chromosome inactivation in cloned mouse embryos. Science *290*,1578-1581.

Elsik, C. G., Tellam, R. L., Worley, K. C., Gibbs, R. A., Muzny, D. M., Weinstock, G. M., Adelson, D. L., Eichler, E. E., Elnitski, L., Guigo, R.,et al.(2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution.Science324, 522–528.

Feng, S., Jacobsen, S.E., Reik, W. (2010). Epigenetic reprogramming in plant and animal development. Science *330*, 622-627.

Fernandez, P.C., Frank, S.R.,Wang, L., Schroeder, M., Liu, S.,Greene, J., Cocito, A., and Amati, B. (2003). Genomic targets of the human c-Myc protein. Genes Dev. *17*, 1115–1129.

Flanagan, Mark E., Blumenkopf, Todd A. et al (2010). Discovery of CP-690,550: A Potent and Selective Janus Kinase (JAK) Inhibitor for the Treatment of Autoimmune Diseases and Organ Transplant Rejection. Journal of Medicinal Chemistry *53* (24), 8468-8484.

Francis et al (2009). Polycomb proteins remain bound to chromatin and DNA during DNA replication *in vitro*. Cell *137* (1), 110-22.

Francis, N.J., Saurin, A.J., Shao, Z., and Kingston, R.E. (2001). Reconstitution of a functional core Polycomb repressive complex. Mol. Cell *8*, 545–556.

Fukada, T., Hibi, M.; Yamanaka, Y., Takahashi-Tezuka, M., Fujitani, Y., Yamaguchi, T., Nakajima, K., Hirano, T. (1996). Two signals are necessary for cell proliferation induced by a cytokine receptor gp130: Involvement of STAT3 in anti-apoptosis. Immunity *5*, 449–460.

Gehring, W. J. (1996). The master control gene for morphogenesis and evolution of the eye. Genes Cells *1,* 11–15.

Gifford et al (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153* (5), 1149-63.

Gotea, V., and Makalowski, W. (2006) Do transposable elements really contribute to proteomes? Trends Genet. 22, 260–267

Gu et al (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477* (7366), 606-10.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A. Wessels, L., de Laat, W., et al. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions.Nature 453, 948–951

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Gurdon et al (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182* (4627), 64-5.

Hamazaki, T., Oka, M., Yamanaka, S., Terada, N. (2004). Aggregation of embryonic stem cells induces Nanog repression and primitive endoderm differentiation. J Cell Sci. Nov 1; 117(Pt 23): 5681-6.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer- blocking activity at the *H19/Igf2* locus. Nature *405*, 486-489.

Hattori, N., Nishino, K., Ko, Y.G., et al (2004). Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. J Biol Chem *279*, 17063–17069.

Hata, K., Okano, M., Lei, H. & Li, E. (2002). DNMT3L cooperates with the DNMT3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983–1993.

Hayashi et al (2007). Germ cell specification in mice. Science *316* (5823), 394-6.

Hayashi, K., Surani, M.A. (2009). Review Resetting the epigenome beyond pluripotency in the germline. Cell Stem Cell. Jun 5; 4(6): 493-8.

Henikoff, S., Furuyama, T., and Ahmad, K. (2004). Histone variants, nucleosome assembly and epigenetic inheritance. Trends Genet. *20*, 320–326.

Herr, W. and Cleary, M.A. (1995). The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. Genes Dev, *9*, 1679-1693.

Hou et a*l* (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science *341*, 651–654.

Humphrey, R.K., Beattie, G.M., Lopez, A.D., Bucay, N., King, C.C., Firpo, M.T., Rose-John, S., and Hayek, A. (2004). Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. Stem Cells *22*, 522–530.

Ichiyanagi, K., Li, Y., Watanabe, T., Ichiyanagi, T., Fukuda, K., Kitayama, J., Yamamoto, Y., Kuramochi Miyagawa, S., Nakano, T., Yabuta, Y., et al. (2011) Locus and domain dependent control of DNA methylation at mouse B1 retrotransposons during male germ cell development.GenomeRes.21, 2058–2066.

llmensee, K. & Hoppe, P. C. (1981). Nuclear transplantation in *Mus musculus*: developmental potential of nuclei from preimplantion embryos. Cell *23*, 9–18.

Iqbal et al (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc Natl Acad Sci USA *108* (9), 3642-7.

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41: 178–186.

Jackson, M. *et al* (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol. Cell. Biol. *24*, 8862–8871.

Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. *33* Suppl, 245–254.

Jagadeeswaran, P., Forget, B. G., and Weissman, S. M. (1981) Short interspersed repetitive DNA elements in eucaryotes: transposable DNA elements generated by reverse transcription of RNA pol III transcripts? Cell 26,141–142

Johnson, Martin H. and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. Cell *24.1*, 71–80.

Jones, P.L., G.J.C. Veenstra, P.A. Wade, D. Vermaak, S.U.Kass, N. Landsberger, J. Strouboulis, and A.P. Wolffe (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genet. *19*, 187–191.

Kagiwada et al (2012). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J, pp. 1-14.

Kahan, B.W. and Ephrussi, B. (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J. Natl. Cancer Inst. *44*, 1015–1036.

Kaneda, M. *et al* (2004). Essential role for *de novo* DNA methyltransferase DNMT3A in paternal and maternal imprinting. Nature *429*, 900–903.

Kang, L., Wang, J., Zhang, Y., Kou, Z., Gao, S. (2009). iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. Cell Stem Cell *5*,135–138.

Karimi et al (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell *8.6*, 676-87.

Kikyo, N., Wade, P. A., Guschin, D., Ge, H. & Wolffe, A. P. (2000). Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. Science *289*, 2360–2362.

Kim, S. H. et al (2004). Differential DNA methylation reprogramming of various repetitive sequences in mouse preimplantation embryos. Biochem. Biophys. Res. Commun. *324*, 58–63.

Kleinsmith, L.J. and Pierce Jr., G.B. (1964). Multipotentiality of single embryonal carcinoma cells. Cancer Res. *24*, 1544–1551.

Kobayashi, H. *et al* (2012). Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet.* 8, e1002440.

Koche et al (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8.1*, 96-105.

Ku, M. *et al* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* 4, e1000242.

Kuff EL, Lueders KK. (1988) The intracisternal A-particle gene family: structure and functional aspects. Adv Cancer Res.;51:183–276

Kunath, T.; Saba-El-Leil, M.K.; Almousailleakh, M.; Wray, J.; Meloche, S.; Smith, A. (2007) FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, *134*, 2895–2902

Kuroda, T., Tada, M., Kubota, H. et al (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol Cell Biol *25*, 2475–2485.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C.,Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860–921.

Labosky, P.A., Barlow, D.P., and Hogan, B.L. (1994). Mouse embryonic germ (EG) cell lines: Transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. Development *120*, 3197–3204.

Leitch et al (2010). Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development *137* (14), 2279-87.

Levasseur, D. N., Wang, J., Dorschner, M. O., Stamatoyannopoulos, J. A. & Orkin, S. H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. Genes Dev*.* *22*, 575–580.

Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. Nature *276*, 565–570.

Li, Mo, Guang-Hui Liu & Juan Carlos Izpisua Belmonte (2012). Navigating the epigenetic landscape of pluripotent stem cells. Nature Reviews Molecular Cell Biology *13*, 524–535.

Li and Sasaki (2011). Genomic imprinting in mammals: its life cycle, molecular mechanisms and reprogramming. Cell Res *21* (3), 466-73.

Li X, Ito M, Zhou F, Youngson N, Zuo X, et al. (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell 15: 547–557

Lienert, F. *et al* (2011)*.* Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet.* 7, e1002090.

Liu, Y., Toh, H., Sasaki, H., Zhang, X., and Cheng, X. (2012). An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev. *26*, 2374–2379.

Lorincz, M.C., Dickerson, D.R., Schmitt, M., and Groudine, M. (2004). Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat. Struct. Mol. Biol. *11*, 1068–1075.

Lubinsky, M.S. & Hall, J.G. (1991) Genomic imprinting, monozygous twinning, and X inactivation. Lancet 337, 1288.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8: A resolution. Nature, 389(6648): 251–260.

Luning Prak ET, Dodson AW, Farkash EA, Kazazian HH Jr: (2003) Tracking an embryonic L1 retrotransposition event. Proc Natl Acad Sci USA, 100 :1832-1837

Lyon MF (1961). Gene Action in the *X*-chromosome of the Mouse (*Mus musculus* L.) *Nature,* 190 (4773): 372–3.

Ma, L. *et al* (2012). Human embryonic stem cell-derived GABA neurons correct locomotion deficits in quinolinic acid-lesioned mice. Cell Stem Cell *10*, 455–464.

Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN, Mager DL. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet. 2:e2

Mancini-DiNardo et al (2003). A differentially methylated region within the gene Kcnq1 functions as an imprinted promoter and silencer. Human Molecular Genetics 12 (3), 283-94.

Mann, J. R., Gadi, I., Harbison, M. L., Abbondanzo, S. J. & Stewart, C. L. (1990). Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: implications for genetic imprinting. *Cell* 62, 251–260.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by terato-carcinoma stem cells. *Proc. Natl. Acad. Sci. 78*, 7634–7638.

Mattick, J.S., (2003). Challenging the dogma: the hidden layer of non-protein- coding RNAs in complex organisms. Bioessays *25*, 930–939.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature *466*, 253–257.

McMahon, S.B., Van Buskirk, H.A., Dugan, K.A., Copeland, T.D., and Cole, M.D. (1998). The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell *94*, 363–374.

Meissner et al. (2009) Derivation and manipulation of murine embryonic stem cells. Methods Mol Biol, vol. 482 pp. 3-19

Meissner, A. *et al* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Meissner (2010). Epigenetic modifications in pluripotent and differentiated cells. Nat Biotechnol *28* (10), 1079-88.

Mikkelsen et al (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454* (7200) pp. 49-55.

Mikkelsen, T. S. *et al* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Mise et al. (2008) Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. Genes Cells vol. 13 (8) pp. 863-77

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. May 30; 113(5): 631-42.

Monk M. (1981). A stem-line model for cellular and chromosomal differentiation in early mouse-development. Differentiation 19(2):71–6.

Moran, J. V., Holmes, S. E., Naas, T. P., DeBerardinis, R. J.,Boeke, J. D., and Kazazian, H. H. Jr. (1996) High frequency retrotransposition in cultured mammalian cells. Cell 87, 917–927

Najm et al. Isolation of epiblast stem cells from preimplantation mouse embryos. Cell Stem Cell (2011) vol. 8 (3) pp. 318-25
Nature Editorial 466, 159 (2010).

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Ohinata et al (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature 436 (7048): 207-13.

Ohshima, K., Hamada, M., Terai, Y., and Okada, N. (1996) The 3′ends of tRNA-derived short interspersedrepetitive elements are derived from the 3′ends of long interspersed repetitive elements.Mol. Cell. Biol. 16,3756–3764

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germ- line competent induced pluripotent stem cells. Nature *448*, 313–317.

Okumura-Nakanishi, S., Saito, M., Niwa, H. et al (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. J Biol Chem *280*, 5307–5317.

Oswald, J. et al (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Otani, J. *et al* (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO Rep. *10*, 1235–1241.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cell. Cell Stem Cell *1*, 299–312.

Pereira, C. F. *et al* (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS Genet*. 4, e1000170.

Phillips, J. E., and Corces, V. G. (2009) CTCF: master weaver of the genome.Cell137,1194–1211

Piccolo et al (2013). Different Roles for Tet1 and Tet2 Proteins in Reprogramming-Mediated Erasure of Imprints Induced by EGC Fusion. Molecular Cell (2013) vol. 49 (6) pp. 1023-1033.

Plagge, A. *et al* (2004). The imprinted signaling protein XLαs is required for postnatal adaptation to feeding. Nature Genet. *36*, 818–826.

Popp et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature (2010) vol. 463 (7284) pp. 1101-5

Ramsahoye, B.H. *et al* (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. USA *97*, 5237–5242.

Rodda, David J., Joon-Lin Chew et al (2005). Transcriptional Regulation of *Nanog* by OCT4 and SOX2. J. Biol. Chem. *280*, 24731-24737.

Roh, T.Y., Cuddapah, S., and Zhao, K. (2005). Active chromatin domains are defined by acetylation islands revealed by genomewide mapping. Genes Dev. *19*, 542–552.

Rowland, B.D., Bernards, R., and Peeper, D.S. (2005). The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. Nat. Cell Biol. *7*, 1074–1082.

Sabour et al. (2011) Identification of genes specific to mouse primordial germ cells through dynamic global gene expression. Human Molecular Genetics vol. 20 (1) pp. 115-25

Saga (2008). Mouse germ cell development during embryogenesis. Curr Opin Genet Dev 18 (4), 337-41.

Saitou, M., Barton, S.C., Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. Nature *418*, 293-300

Santos, F., Hendrich, B., Reik, W. & Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev. Biol*. 241*, 172–182.

Saxonov, S., Berg, P. & Brutlag, D. L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl Acad. Sci. USA 103, 1412–1417.

Schneuwly, S., Klemenz, R. & Gehring, W. J. (1987). Redesigning the body plan of *Drosophila* by ectopic expression of the homoeotic gene *Antennapedia*. Nature *325*, 816–818.

Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooper- berg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genomewide chromatin analysis of a higher eukaryote. Genes Dev. *18*, 1263–1271.

Seki, Y. *et al* (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev. Biol. *278*, 440–458.

Senner CE, Brockdorff N (2009). Xist gene regulation at the onset of X inactivation=genetics-development. Current opinion in genetics & development *19* (2), 122–26.
*S*hapiro*, J. A.* (2010). *Mobile DNA and evolution in the 21st century.* Mob. DNA 1: 4*.*

*S*hapiro*, J. A.* (1969). *Mutations caused by the insertion of genetic material into the galactose operon of Escherichia coli.* J. Mol. Biol. *40*, 93*–105.*

Shiba, Y. *et al* (2012). Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. Nature 489, 322–325.

Silva, J., Chambers, I., Pollard, S. & Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature *441*, 997–1001.

Simon, J. A. & Kingston, R. E. (2009). Mechanisms of Polycomb gene silencing: knowns and unknowns. Nature Rev. Mol. Cell Biol. *10*, 697–708.

Simonsson, S., Gurdon, J. (2004). DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol *6*, 984–990.

Smallwood, S. A. *et al* (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nature Genet. *43*, 811–814.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Somers, A. *et al* (2010). Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette Stem Cells *28*, 1728–1740.

Sorek, R., Ast, G., and Graur, D. (2002)Alu containing exons are alternatively spliced.Genome Res.12,1060–1067.

Stadtfeld, M. *et al* (2008). Induced pluripotent stem cells generated without viral integration Science, *322*, 945–949.

Stadtfeld, M. and K. Hochedlinger (2010). Induced pluripotency: history, mechanisms, and applications Genes Dev. *24*, 2239–2263.

Stevens, L.C. and Little, C.C. (1954). Spontaneous testicular tera-tomas in an inbred strain of mice. Proc. Natl. Acad. Sci. *40*, 1080–1087.

Strahl, B. D., Allis, C. D. (2000). The language of covalent histone modi!cations. Nature *403*(6765), 41–45.

Surani, M. Azim Katsuhiko Hayashi, and Petra Hajkova (2010). Genetic and Epigenetic Regulators of Pluripotency DOI 10.1016/j.cell.2007.02.010.

Tada et al (1998). Epigenotype switching of imprintable loci in embryonic germ cells. Dev Genes Evol *207* (8), 551-61.

Tada et al (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. EMBO J *16* (21), 6510-20.

Tada et al (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11* (19),1553-8.

Tahiliani et al (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324* (5929), 930-5.

Tahsin, Stefan Barakat and Joost Gribnau (2010). X Chromosome Inactivation and Embryonic Stem Cells. *The Cell Biology of Stem Cells* edited by Eran Meshorer and Kathrin Plath. Landes Bioscience and Springer Science+Business Media.

N. Takagi, K. Abe (1990) Detrimental effects of two active X chromosomes on early mouse development Development, 109 pp. 189–201

Takahashi and Yamanaka (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126* (4), 663-76.

Takahashi et al (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131* (5), 861-72.

Takaoka, Katsuyoshi and Hiroshi Hamada (2012). Cell fate decisions and axis determination in the early mouse embryo. Development *139,* 3-14.

Tanaka, Y., Yamashita, R., Suzuki, Y., and Nakai, K. (2010) Effects of Alu elements on global nucleosome positioning in the human genome.BMC Genomics 11,309.

Tesar. P.J., Chenoweth, J.G., Brook, F.A. et al (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Thomson et al (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464* (7291), 1082-1086.

Thorvaldsen, J. L., Duran, K. L. and Bartolomei, M. S. (1998). Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. Genes Dev. *12*, 3693-3702.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K. and Niwa, H.(2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909-918

Tsumura, A. *et al* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells *11*, 805–814.

Tsunoda, Y. & Kato, Y. (1998). Not only inner cell mass cell nuclei but also trophectoderm nuclei of mouse blastocysts have a developmental totipotency. J. Reprod. Fertil*.* *113*, 181–184.

Uematsu, A.; Yorifuji, T.; Muroi, J.; Kawai, M.; Mamada, M.; Kaji, M.; Yamanaka, C.; Momoi, T.; Nakahata, T. (2002). "Parental origin of normal X chromosomes in Turner syndrome patients with various karyotypes: Implications for the mechanism leading to generation of a 45,X karyotype". *American Journal of Medical Genetics* 111 (2): 134–139.

Urness LD, Li C, Wang X, Mansour SL (2008) Expression of ERK signaling inhibitors Dusp6, Dusp7, and Dusp9 during mouse ear development. Dev Dyn 237: 163–169

Vallier, Ludovic, Sasha Mendjan, Stephanie Brown, Zhenzhi Chng, Adrian Teo, Lucy E. Smithers, Matthew W. B. Trotter, Candy H.-H. Cho, Amelie Martinez, Peter Rugg-Gunn, Gabrielle Brons, Roger A. Pedersen (2009). Activin/Nodal signaling maintains pluripotency by controlling Nanog expression. Development April 15; 136(8): 1339–1349.

Veitia RA, Bottani S, Birchler JA (2013) Gene dosage effects: Nonlinearities, genetic interactions, and dosage compensation. Trends Genet 29(7):385–393.

Voo, Shin K., Carlone, D. L., Jacobsen, B. M., Flodin, A. & Skalnik, D. G. (2000). Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. Mol. Cell. Biol. *20*, 2108–2121.

Waddington, C.H. (1942). The epigenotype. Endeavour *1*, 18–20.

Wallis MC, Waters PD, Graves JA (2008). "Sex determination in mammals - Before and after the evolution of SRY". Cell. Mol. Life Sci. 65 (20): 3182–95.

Wang, S. *et al* (2013). Human iPSC-derived oligodendrocyte progenitor cells can myelinate and rescue a mouse model of congenital hypomyelination. Cell Stem Cell *12*, 252–264.

Wang, L., Brown, J.L., Cao, R., Zhang, Y., Kassis, J.A., and Jones, R.S. (2004). Hierarchical recruitment of polycomb group silencing complexes. Mol. Cell *14*, 637–646.

Wang, G., Zhang, H., Zhao, Y., Li, J., Cai, J., Wang, P., Meng, S., Feng, J., Miao, C., Ding, M., et al. (2005). Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. Biochem. Biophys. Res. Commun. *330*, 934–942.

L. Warren *et al* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA Cell Stem Cell *7*, 618–630.

Weber, M. et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nature Genet. 37, 853–862.

Willadsen, S. M. (1986). Nuclear transplantation in sheep embryos. Nature *320*, 63–65.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. S. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810–813.

Woltjen, K. *et al* (2009). PiggyBac transposition reprograms fibroblasts to induced pluripotent stem cells Nature *458*, 766–770.

Wu, C. T., and Morris, J.R. (2001). Genes, genetics, and epigenetics: a correspondence. Science *293*, 1103–1105.

Wu and Zhang (2010). Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol *11* (9), 607-20.

Wu and Zhang. (2014)Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell vol. 156 (1-2) pp. 45-68

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat. Biotechnol*.* *20*, 1261–1264.

Yang et al (2010). Regulation beyond genome sequences: DNA and histone methylation in embryonic stem cells. Front. Biol. China, pp. 1-7.

Ying, Qi-Long, Jennifer Nichols, Ian Chambers,and Austin Smith (31 Oct. 2003). BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. Cell *115*, 281–292.

Ying, Qi-Long, Jason Wray, Laura Batlle-Morera, Bradley Doble,James Woodgett, Philip Cohen & Austin Smith (22 May 2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-23.

Yoshimizu, T., Obinata, M., Matsui, Y. (Feb 2001). Stage-specific tissue and cell interactions play key roles in mouse germ cell specification. Development *128*(4), 481-90.

Yu, J., Angelin-Duclos, C., Greenwood, J., Calame, K. (April 2000). Transcriptional repression by blimp-1 (PRDI-BFI) involves recruitment of histone deacetylase. Mol Cell Biol. 20(7), 2592-603.

Yu, F. *et al* (2011). Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion. Oncogene *30*, 2161–2172.

Zaehres, H., Lensch, M. W., Daheron, L. et al (2005). High-efficiency RNA interference in human embryonic stem cells. Stem Cells *23*, 299–305.

Zhang Y, Maksakova IA, Gagnier L, van de Lagemaat LN, Mager DL. (2008) Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements. PLoS Genet. 4:e1000007

Zhou, H. *et al* (2009). Generation of induced pluripotent stem cells using recombinant proteins Cell Stem Cell *4*, 381–384.

Zhou et al (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet *12* (1), 7-18.

SUPPLEMENTAL

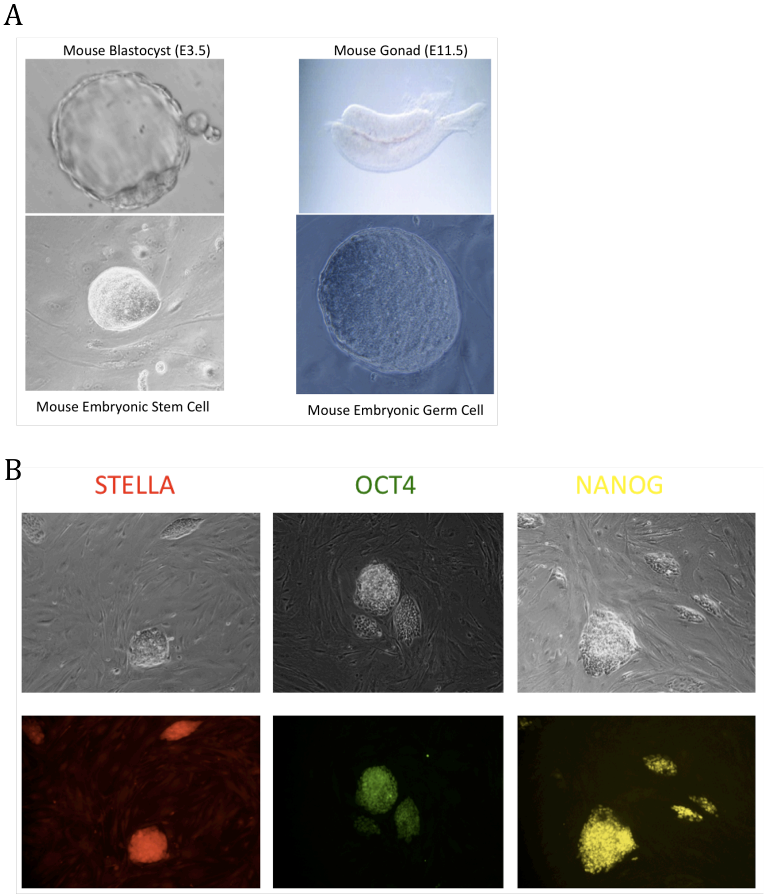
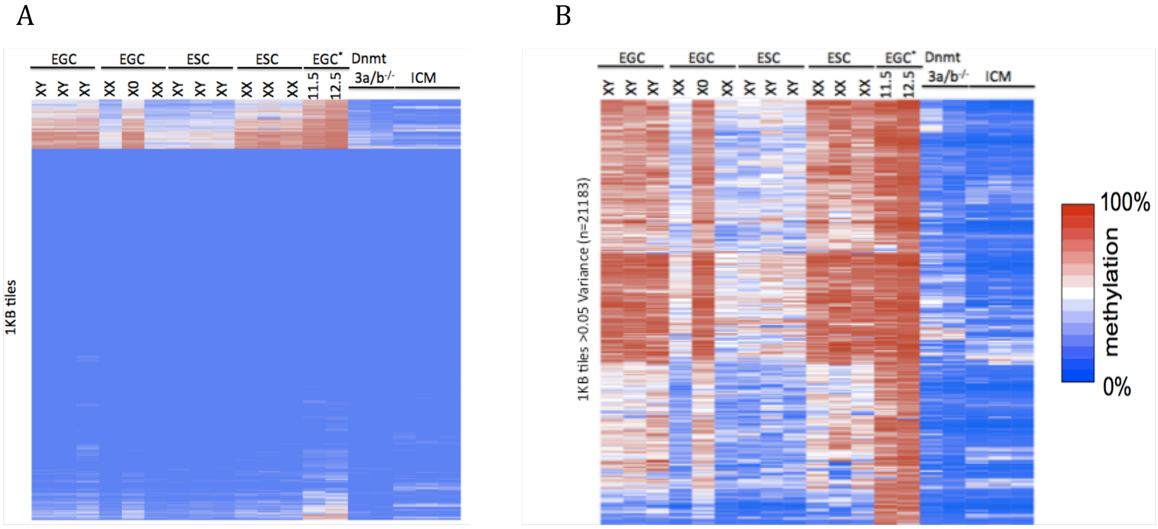


Figure S1. Phenotypic and genotypic similarities between EGCs and ESCs. (A) Different developmental stages for the isolation of ESCs and EGCs. The mouse bastocyst is isolated at embryonic day 3.5. Explantation results in ESCs with a rounded morphology. EGCs are isolated from the PGCs present in the gonads of mice from embryonic day 11.5-12.5. Explantation results in EGCs with a similar rounded morphology to ESCs. (B) EGCs express the same key pluripotent factors present in ESCs. Immunostaining shows the presence of stella, oct4 and nanog in EGCs.



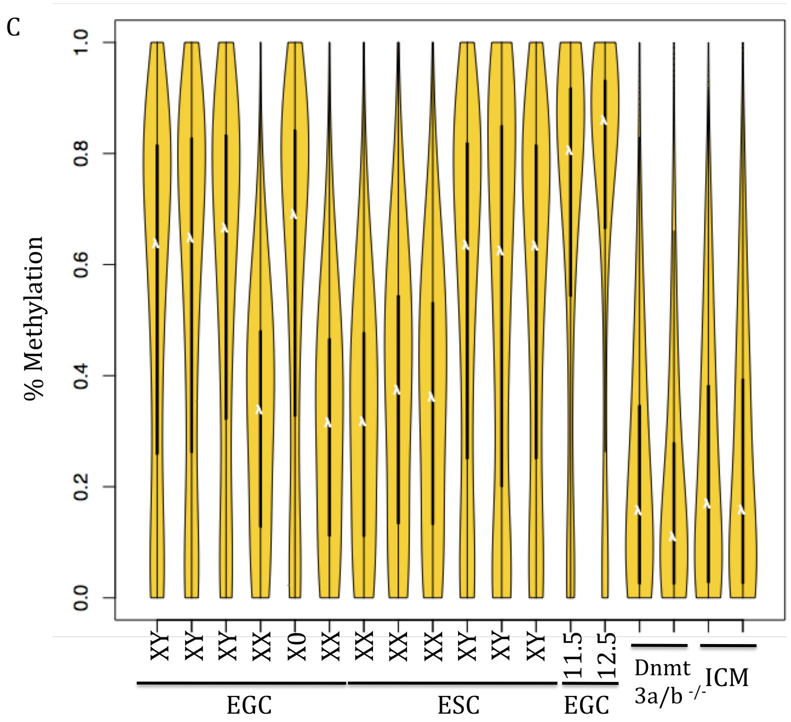


Figure S2. Whole genome methylation analysis of EGCs and ESCs.(A) 1 KB tiles of the entire genome for each individual cell line. (B) 1 KB tiles showing a variance >0.05 for each individual cell line. (C) Violin plots for whole genome 1KB tiling methylation levels. EGCs 11.5 and 12.5 are XY.

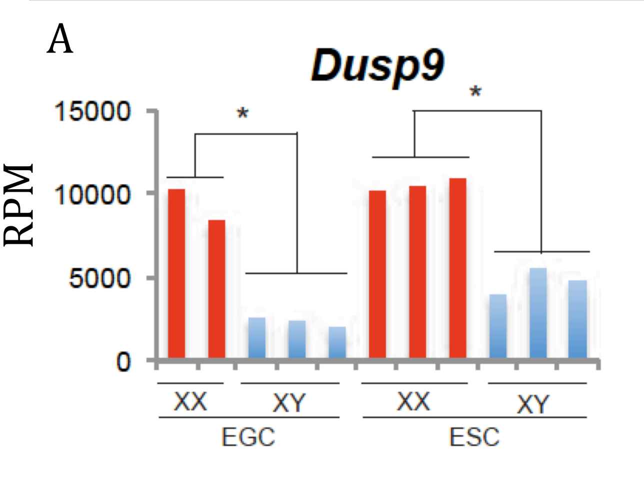
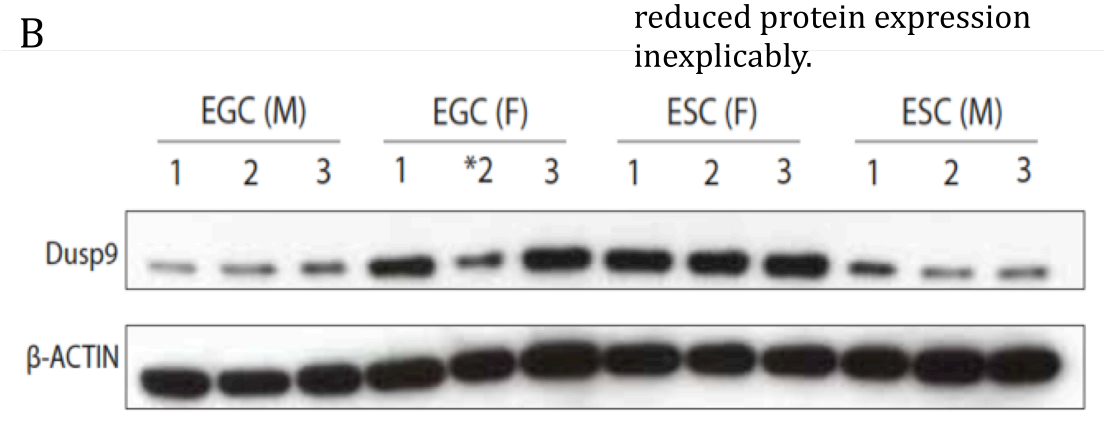


Figure S4. Differential Dusp9 expression in male and female cells. (A) RNA-seq results (reads per million) for Dusp9 expression in EGCs and ESCs. Dusp9 expression is shown to be higher in female cells compared to males (P=0.5). (B) Western blot analysis shows higher levels of Dusp9 protein in female cells. EGC (F) sample #2 shows reduced protein expression inexplicably.



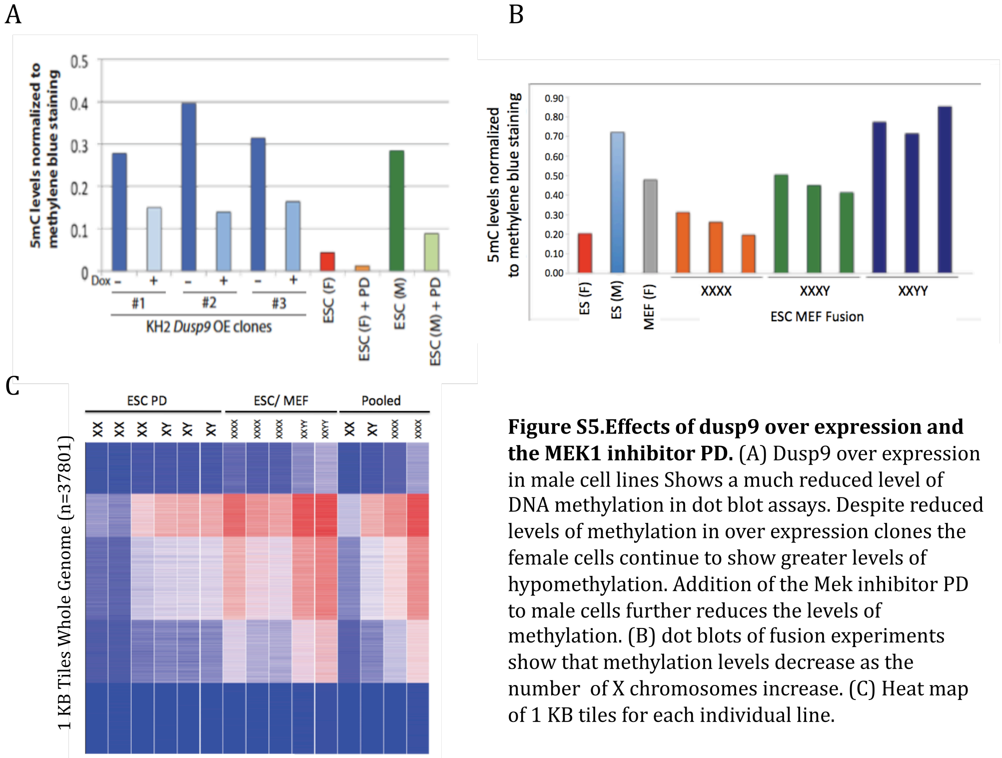
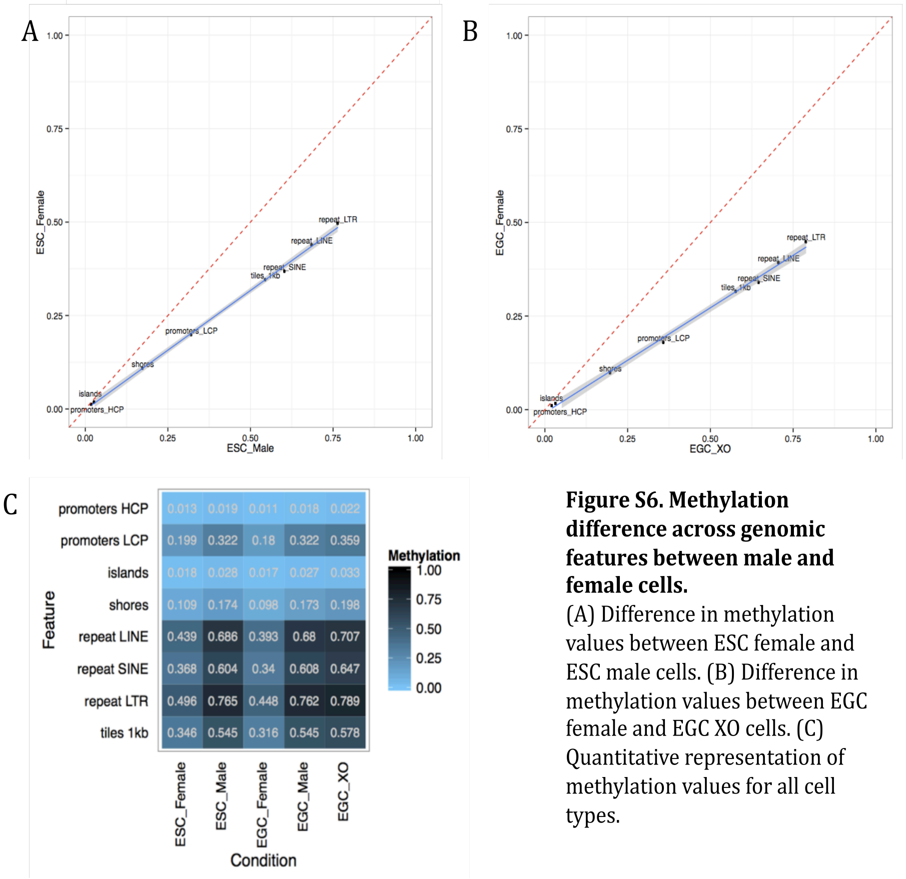


Figure S5.Effects of dusp9 over expression and the MEK1 inhibitor PD. (A) Dusp9 over expression in male cell lines Shows a much reduced level of DNA methylation in dot blot assays. Despite reduced levels of methylation in over expression clones the female cells continue to show greater levels of hypomethylation. Addition of the Mek inhibitor PD to male cells further reduces the levels of methylation. (B) dot blots of fusion experiments show that methylation levels decrease as the number of X chromosomes increase. (C) Heat map of 1 KB tiles for each individual line.



Feature	ESC_Female	ESC_Male	EGC_Female	EGC_Male	EGC_XO
promoters HCP	0.013	0.019	0.011	0.018	0.022
promoters LCP	0.199	0.322	0.18	0.322	0.359
islands	0.018	0.028	0.017	0.027	0.033
shores	0.109	0.174	0.098	0.173	0.198
repeat LINE	0.439	0.686	0.393	0.68	0.707
repeat SINE	0.368	0.604	0.34	0.608	0.647
repeat LTR	0.496	0.765	0.448	0.762	0.789
tiles 1kb	0.346	0.545	0.316	0.545	0.578

Figure S6. Methylation difference across genomic features between male and female cells.

(A) Difference in methylation values between ESC female and ESC male cells. (B) Difference in methylation values between EGC female and EGC XO cells. (C) Quantitative representation of methylation values for all cell types.

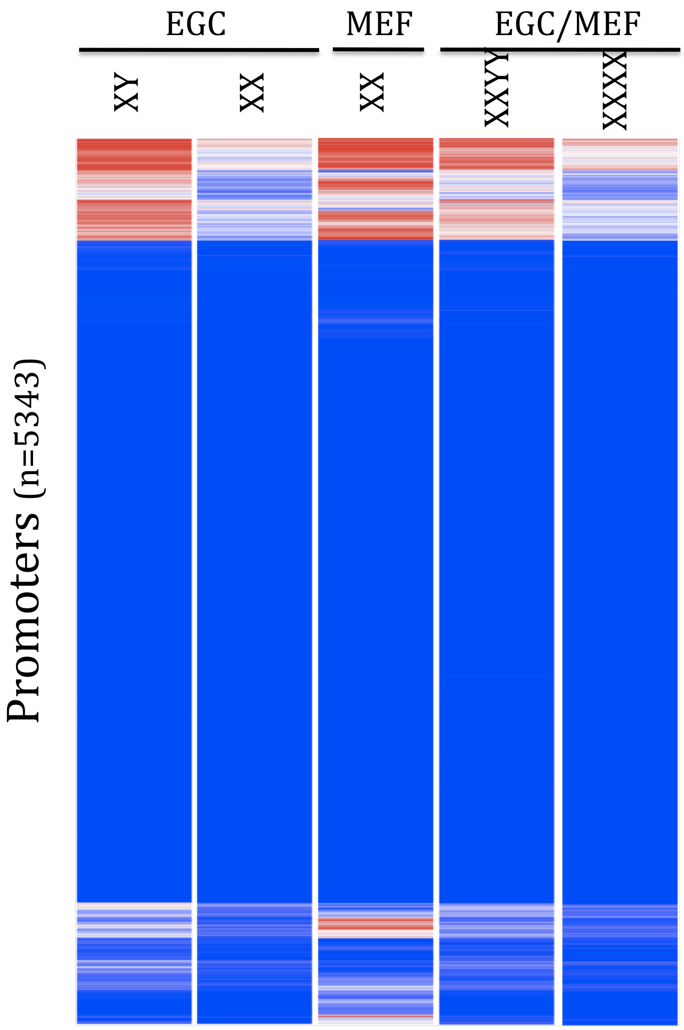


Figure S7. Methylation differences across all promoters in EGCs, MEFs, and fused cells. Female EGCs increase hypomethylation in fusion samples. Fused male EGCs exhibit a level of methylation lower than that of individual Male EGCs and MEFs. This may be a result of an increased dosage of Dusp9 from the second X chromosome.